Characterisation of a sugar beet gene encoding a novel *cdk*-like kinase

by

Ana Ivanova Atanassova

Thesis submitted for the degree of
Doctor of Philosophy

The Norman Borlaug Institute for Plant Science Research

De Montfort University

Leicester

September 2002

This thesis is enterely my own work and has not been submitted for another degree.

Date

06.10.2013

A.I. Atanassova

ABSTRACT

The cyclin-dependent kinases (CDKs) are major regulators of the cell cycle progression in eukaryotes. Advancement in the cell cycle research has shown that the underlying mechanisms of cell cycle controls in animals and plants are very similar. Nevertheless, plants possess specific controls, which accommodate for their unique developmental programs and body patterns. Numerous CDKs and CKD-related genes have been identified in plants, and recently they have been classified into several plant specific subgroups.

The present work describes the isolation and further characterisation of a novel CDK-related gene from sugar beet, *Bvcrk1*. Previously, it was demonstrated that the transcripts of the *Bvcrk1* are rapidly induced at the G0 to G1/S transition in suspension cultured sugar beet cells, and here it is demonstrated that transcripts of the gene are localised to primary and secondary meristems. Phylogenetic analysis of *Bvcrk1* indicated that although it has many features in common with the CDK family members, it can not be classified as a true CDK due to important changes in the kinase domain and long N- and C-terminal extensions. The unique PSTAIRE signature, namely ESV(K/R)FMARE, within the kinase domain was found to be specific to plants. The *Bvcrk1* appears to be the first gene form this new CDK-related subfamily of protein kinases to be characterised in more detail, and we propose that it may be involved in cell cycle or early differentiation events specific to plants.

CONTENTS

CHAPTER 1 1

1.1 INTRODUCTION 1

1.2 CELL DIVISION CYCLE AND ITS CONTROLS 3

1.2.1 INTRODUCTION 3
1.2.2 CYCLIN-DEPENDENT PROTEIN KINASES 4
1.2.2.1 Basic structural characteristics of the CDKs 5
1.2.2.2 The role of cyclin binding 6
1.2.2.3 Regulation of the CDK activity by phosphorylation 7
1.2.2.4 CDK inhibitors 8
1.3 PLANT CDKS AND REGULATORS OF THE PLANT CELL CYCLE 8
1.3.1 Overview of plant CDKs 9
1.4 THE G1/S AND G2/M PHASE TRANSITIONS AND REGULATORS OF THE PLANT CELL CYCLE 13
1.4.1 The G1/S controls 13
1.4.2 G2/M transition in plant cells 18

CHAPTER 2 PCR BASED ISOLATION OF CYCLIN-LIKE SEQUENCES 22

2.1 MATERIALS 22
2.1.1 Plant material 22
2.1.2 Bacterial strains and cloning vectors 22
2.1.3 PCR primers 22
2.1.4 Reagents, media, solutions, enzymes and kits used 23
2.2 METHODS 23
2.2.1 Germination of seedlings and collection of samples for mRNA isolation 23
2.2.2 mRNA isolation 23
2.2.3 Synthesis of cDNA 24
2.2.4 PCR amplification of sugar beet cyclin -like sequences 25
2.2.5 Cloning of PCR fragments 26
2.3 RESULTS 27
2.3.1 PCR based isolation of cyclin sequences 27

CHAPTER 3 GENOMIC LIBRARY CONSTRUCTION AND SCREENING 30

3.1 MATERIALS 30
3.1.1 Plant material 30
3.1.2 Cloning vectors and bacterial strains 30
3.1.3 Kits used for the construction and screening of the genomic library 30
3.1.4 Restriction and modifying enzymes, media and solutions 31
3.2 GENOMIC LIBRARY CONSTRUCTION 31
3.2.1 Plant maintenance and collection of samples for DNA isolation 31
3.2.2 Plant DNA isolation 32
3.2.3 Partial digest of genomic DNA 33
3.2.4 Partial fill-in of Sau 3A I digested DNA 35

3.2.5 Ligation of the partially filled-in DNA fragments to Lambda FIX II vector arms 35
3.2.6 In vitro packaging of the recombinant lambda DNA and titration of the genomic library 36
3.2.7 Estimation of the average size of the cloned genomic inserts and the quality of the genomic libraries 37
3.2.8 Amplification of the genomic library 38
3.2.9 Lambda phage amplification 39
3.2.10 Lambda DNA isolation 40
3.3 SUGAR BEET GENOMIC LIBRARY SCREENING 41
3.3.1 Probes used for the screening of the library 41
3.3.2 Maintenance of the clones and probes labelling 42
3.3.3 Plaque lifts 42
3.3.4 Hybridisation 42
3.3.5 Screening results 43

CHAPTER 4 THE *BVCRK1* GENE: SEQUENCE, INTRON-EXON ORGANISATION AND THE PREDICTED CRK1 PROTEIN 45

4.1 MATERIALS 45
4.1.1 Cloning vectors and bacterial strains 45
4.1.2 Primers 45
4.1.3 Kits and special reagents 45
4.1.4 Restriction and modifying enzymes, media and solutions 45
4.2 METHODS 46
4.2.1 Characterisation of the crk1 positive plaques 46
4.3 COPY DNA ISOLATION 54
4.3.1 Isolation of the 5'-end of the crk1 cDNA 54
4.3.2 Messenger RNA isolation 55
4.3.3 Reverse transcription 56
4.3.4 Ligation of the single stranded cDNA 56
4.3.5 PCR amplification of the 5'-end of the crk1 gene 56
4.3.6 Isolation of the 3'-end of the crk1 cDNA 59
4.4 DNA SEQUENCE ANALYSIS 62
4.5 RESULTS 63
4.5.1 Genomic DNA of the Bvcrk1 gene 63
4.5.2 5'-cDNA extension 69
4.5.3 3'-cDNA extension 72
4.5.4 Full-length cDNA of the crk1 gene 73
4.5.5 Analysis of the predicted CRK1 peptide 77

CHAPTER 5 EXPRESSION OF *BVCRK1* GENE 87

5.1 MATERIALS AND METHODS 87
5.1.1 Plant material 87
5.1.2 Preparation of the plant material 88
5.1.3 Preparation of slides 89
5.1.4 Probes for in situ hybridisation 90

- *Probes used in the hybridisations* 90
- *Probe preparation* 90
- *Probe labelling by in vitro transcription* 90

5.2. IN SITU HYBRIDISATION 91

5.2. 1 Prehybridisation procedures 91

5.2.2 Hybridisation 92

5.2.3 Signal detection 93

5.2.4 Microscopy and photography 94

5.3. RESULTS 94

CHAPTER 6 DISCUSSION **104**

6.1 THE PLANT CELL CYCLE 104

6.2 THE SUGAR BEET CELL CYCLE AND MANIPULATION OF STORAGE ROOT DEVELOPMENT 104

6.3 ISOLATION AND CHARACTERISATION OF A NOVEL CDC2-LIKE GENE FROM SUGAR BEET 105

6.4 THE PREDICTED BVCRK1 PROTEIN 106

6.5 POSSIBLE FUNCTIONS OF THE CRK1 PROTEIN FAMILY 107

6.6 FUTURE WORK 107

APPENDIX **109**

MEDIA AND SOLUTIONS 109

LIST OF PRIMERS USED FOR THE ISOLATION AND SEQUENCING OF THE BVCRK1 CDNA 110

LIST OF PRIMERS USED FOR THE SEQUENCING OF THE GENOMIC CLONES OF BVCRK1 GENE 111

LIST OF ABBREVIATIONS **112**

LIST OF REFERENCES **113**

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr Adrian Slater for reading through the manuscript and giving me helpful advice and comments.

I thank full-heatedly Jenny, Bill and John not only for their support and practical advice in the lab, but also for their great company and friendship.

I am grateful to Vicky, and her family, for making me very welcome in their home.

I would like to thank also my parents for their hope in me, ...and for their truly strong wish for me to finish this work.

I am grateful to the De Montfort University for providing my research grant.

I express my gratefulness to Jorge for his love, support and great sense of humour and to Clara and Francesco for being the most friendly and smiley of children, even in hard times.

CHAPTER 1 BACKGROUND AND LITERATURE REVIEW

1.1 BACKGROUND OF THE PRESENTED WORK

The "low environmental impact-high yield sugar beet" concept was introduced by the team of the De Montfort University Norman Bourlaug Institute in recognition that the classical breeding techniques have improved the crop mainly in terms of its sugar storage capacity paying little attention to the crops environmental impact. Indeed, sugar beet storage root accumulates sucrose to levels of 18-20% of its fresh weight. In the EC, two million hectares of farmland are given over to sugar beet growing. Its harvesting, consequent cleaning and transportation mean that yearly as much as 3 million tonnes of soil have to be removed from the sugar beet at the processing factories which brings high financial and environmental costs due to the loss of fertile soil, water, fuel and energy.

A strategy for production of a low environmental impact sugar beet has been developed by the NBI (Elliott and Weston, 1993; Thomas *et al.*, 1993) which could be presented in two parts. On one hand, it was hypothesised that via genetic manipulation of the table beet cultivar of *Beta vulgaris* one can produce a plant with the sugar accumulation capacity and juice purity of the best sugar beet cultivars. The argument goes that the table beet has the desired shape of the root (globular) and growth habit (it grows high in the soil so it is easy to harvest and retains less soil on its surface) but for high sucrose and high juice purity certain structural modifications are necessary. It was proposed that the existing outer cambia of the table beet have to be activated for cell division so that a larger root body is produced with more than twice as many sucrose storage cells available (Elliott *et al.*, 1993). On the other hand, the genetic improvement of the shape of the tap root of sugar beet holds further potential for reducing the environmental costs of sugar production.

The implementation of the above mentioned goals requires a detailed knowledge about the intricate developmental program of the beet, with understanding about many physiological and morphological events occurring from the emergence of the seedling until the formation of the storage root. The focus of attention of the NBI team was placed on studying processes of cell division, enlargement and differentiation of the sugar beet root.

At the beginning, it was observed that during development of the sugar beet root changes in the level of phytohormones correlate with the establishment and differentiation of the storage organ (Elliot *et al.*, 1993). To investigate further the link between cell division and hormones, a cell suspension culture of sugar beet was used to study the correlation of the presence and activity of different hormones in the media and the divisional activity of the cells (Fowler *et*

al., 1998). The suspension culture system has enabled the first description of the transcriptional activities of several cell cycle related genes such as a putative *cdc2* homologue, *Bvcdc2*, a A-type cyclin gene, Betvu;*CycA2*, and a CDK-related gene *Bvcrk1*.
At the time of the isolation of the *Bvcrk1* partial sequence, it was obvious that no homologues existed in the databank, and that the gene showed cell cycle dependent expression pattern in the sugar beet suspension system. Due to the gene uniqueness, it was decided to isolate the full cDNA and genomic sequence and to learn more about its function in the sugar beet root. A big part of the work on the further characterisation of the *Bvcrk1* is presented in the following chapters of this thesis. To date, we can state that based on the sequence data, *Bvcrk1* belongs to the group of Serine-Threonine cyclin-dependent protein kinases (CDKs) but that it is a plant-specific protein as no homologues exist in other organisms. Furthermore, our expression studies show that the *Bvcrk1* gene has abundant but restricted to the meristems expression pattern. Thus, we propose that *Bvcrk1* might function in the regulation of some specific aspects of the plant cell cycle.

1.2 CELL DIVISION CYCLE AND ITS CONTROLS

1.2.1 Introduction

No matter how small or big a living organism is, its life cycle begins with one event – cell division, a process that creates great numbers of individual cells from unicellular organisms, or builds up the bodies of more complex life forms. In multicellular organisms, cell multiplication is an integral part of the organism's growth and development, and this is reflected in a plethora of regulatory and control mechanisms that are in place to ensure the correct and "smooth" functioning of the organism in response to a variety of internal and external signals. The picture of how developmental and environmental signals are transmitted to the cell cycle machinery is still far from complete but the basic aspects of the cell division cycle have been extensively researched for the last twenty to thirty years. In order to reproduce, cells must replicate their chromosomes and then segregate them into the new daughter cells, which will divide in their turn. The execution of such a cycle must be under rigorous controls in order to ensure that the genetic material is correctly duplicated, and segregated equally.

The cell division cycle is divided into four discrete phases S (DNA synthesis), preceded by a gap G1 and followed by a gap G2 before M phase (mitosis). The DNA replication is preceded by two events: activation of the origins of replication, the regions of the chromosome where DNA replication initiates, and activation of the enzymes needed for the DNA synthesis. In M phase the replicated chromosomes segregate into two new cells.

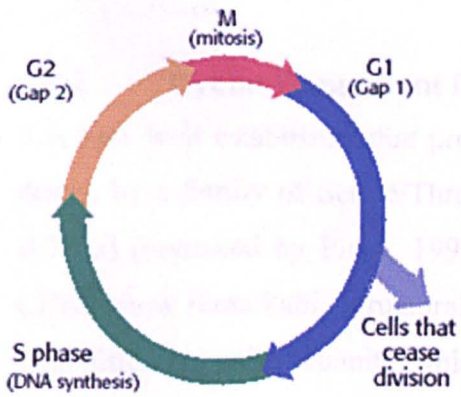


Figure 1.1 The four cell cycle stages

The G1, S, G2 and M- phases of the cell division cycle are depicted in the diagram to demonstrate the continuity of the process. Cells may exit from the cycle (indicated by an arrow) normally before a new round of DNA replication. Plants can exit the cell cycle in G2 as well (not indicated).

The diagram is copied from the web site http://www.biology.arizona.edu.

A great deal of knowledge about the regulation of division has emerged from the study of different model organisms such as yeast, amphibians, *Drosophila*, cultured mammalian cells and plants. A unified model of the cell cycle has emerged, where cells pass through consecutive cell cycle phases, from G1 to S to G2 to M (figure 1.1). Orderly progression through the cell cycle is safeguarded by a series of positive and negative feedback mechanisms, which operate at a variety of control points or checkpoints (Nurse, 1990; 2000). Essentially, three major control points can be described: in the late G1, at the G2/M phase boundary and at the metaphase to anaphase transition; The point in G1 is know as START in yeast (Nurse, 1990) and the restriction point (R) in mammals (Pardee, 1989) and is the point past which the cell can no longer undergo any developmental fate other than DNA replication. Later in the cycle, at the G2/M boundary, cells pass through a point beyond which entry into mitosis can no longer be reversed. Exit from mitosis begins with the separation of replicated chromosomes at a checkpoint operating at the metaphase to anaphase transition (Pines and Rieder, 2001). Once past this point and following cytokinesis, conditions are established to prepare the cell for thenext round of DNA replication.

However, the 'normal' cell cycle order of events with the typical G1-S-G2-M phases does not always hold true. In some cases, the dependence of S-phase on M-phase completion is bypassed, resulting in numerous rounds of DNA replication without division (Nagl, 1995). Another example is the early embryonic divisions in *Drosophyla* and in some amphibians which lack the gap phases. Cells may also enter and remain in state of quiescence (G0) with or without the capacity to resume cycling (Warner-Washburne et al., 1993).

1.2.2 Cyclin-Dependent Protein Kinases

It is now well established that progression through the major transitions of the cell cycle is driven by a family of Serine/Threonine kinases known as cyclin dependent protein kinases (CDKs) (reviewed by Pines, 1993; Sherr, 1993). Throughout the eukaryotic kingdom, the CDKs show remarkable structural and functional conservation. The CDKs form a complex with different cyclin subunits, which render the complex fully active. Apart from ensuring the activation of the catalytic kinase subunit at precise time points, the cyclins also contribute to the correct substrate targeting and subcellular localisation of the complex (Morgan, 1997; Brown *et al.*, 1999; Kong *et al.*, 2000). Further modulation of the activity of the CDK/cyclin complexes is provided by reversible phosphorylation/dephosphorylation, and binding of inhibitory proteins (Pines, 1996).

1.2.2.1 Basic structural characteristics of the CDKs

The cyclin-dependent protein kinases belong to the family of Serine/Threonine protein kinases (Hanks et al., 1988; Hanks and Hunter, 1995) and are characterised by the presence of a highly conserved central catalytic core (the kinase domain), flanked by N- and C- terminal extensions of different size. Apart from the ATP- and substrate-binding sites, the CDKs contain in their catalytic core several highly conserved regulatory phosphorylation sites (figure 1.2). One major helix, the C (PSTAIRE) helix is present in the N-terminal lobe of the kinase domain and has been shown to be very important for the association of the cyclin subunit (Jeffrey et al., 1995).

The canonical PSTAIRE motif in the C helix is present in the yeast *cdc2*/CDC28 and human CDK1-3. It has been used as a reference in the structural classification of the different CDKs isolated from animals (reviewed by Nigg, 1995), and more recently from plants (Mironov et al., 1999; Joubès et al., 2000). A summary of the information about human CDKs is presented in table 1.1

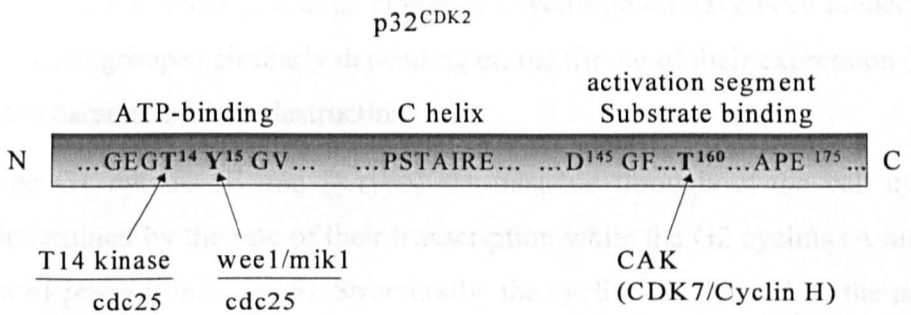


Figure 1.2 Major regulatory and phosphorylation sites in human Cyclin-dependent kinase 2 (CDK2). The schematic drawing illustrates the catalytic core of CDK2, with the major regulatory and phosphorylation sites, which are conserved in all known CDKs. The CDK is negatively regulated by phosphorylation on Thr14 and Tyr15 by Wee1/Mik1 kinase activity. The same residues are dephosphorylated by the dual-specificity phosphatase Cdc25. The phosphorylation of Thr160 in the T-loop by the Cdk-activating kinase (CAK) renders the CDK/cyclin complex fully active. The D^{145}GF...to...APE^{172} segment at the carboxyl-terminal lobe denotes the region involved in the base catalysed transfer of the γ-phosphate of the ATP to the substrate. The PSTAIRE motif in the C-helix in subdomain III is involved in binding to cyclins and confers some level of specificity as to which CDK binds to which cyclin.
The drawing of the kinase domain and the regions indicated within it is not to scale. Composed from DeBondt et al., 1993; Jeffrey et al., 1995; Nigg, 1995; Brown et al., 1999.

Table 1.1 **Human CDK subunits and summary of their proposed function**

Name	Principal regulatory subunits*	Function	PSTAIRE motif
CDK1	Cyclins A, B1, B2, B3	G2/M transition	PSTAIRE
CDK2	Cyclins A, E, D1, (D2, D3)	G1/S transition and S	PSTAIRE
CDK3	p70 (ik3-1,interactor-1with cdk3)	G1/S transition	PSTAIRE
CDK4	Cyclins D1 (D2?), D3	G1/S transition	PISTVRE
CDK5	p35	neurofilament phosphorylation	PSSALRE
CDK6	Cyclins D1 (D2, D3?)	G0/G1 and G1/S transitions	PLISTRE
CDK7	Cyclin H, MAT-1	activates CDK1-4 and CDK6	NRTALRE
CDK8	cyclin C	G1/S transition	SMSACRE
CDK9	Cyclins T1, T2	transcription elongation, not cell cycle regulated	PITALRE
CDK10	unknown	G2/M transition?	PISSLRE

1.2.2.2 The role of cyclin binding

In monomeric form, the kinases are inactive and require cyclin binding for catalytic activity. Cyclins were first identified in sea urchin eggs as proteins which were synthesised continuously during interphase and underwent rapid destruction after each mitosis (Evans et al., 1983). Currently, a large number of cyclin genes have been cloned and characterised and they are grouped similarly depending on the timing of their expression (G1 or G2 phase) or by the character of their destruction.

The G1 cyclins (E and D type) are unstable throughout the cell cycle and their level is determined by the rate of their transcription while the G2 cyclins (A and B) are unstable only in M phase (Pines, 1996). Structurally, the cyclins are defined by the presence of a conserved 100 amino acid region, the cyclin box, which takes part in the CDK binding and activation (Lees and Harlow, 1993; Morgan, 1995). The degradation of G2 cyclins begins at the end of metaphase (Cyclin A) to anaphase exit (Cyclin B) and causes the inactivation of the CDK. The degradation of G2 cyclins is mediated by a small region at their N-terminal part called the destruction box. The destruction box targets the cyclin subunit for proteolytic degradation by the 26S proteasome via the anaphase-promoting complex/cyclosome (APC/C) and ubiquitination (Zachariae and Nasmyth, 1999).

* The data in table 2.1 is collected from the following references: Nigg, (1995); Matsuoka et al., (2000); deFalco et al., (1998); Sergere et al., (2000).

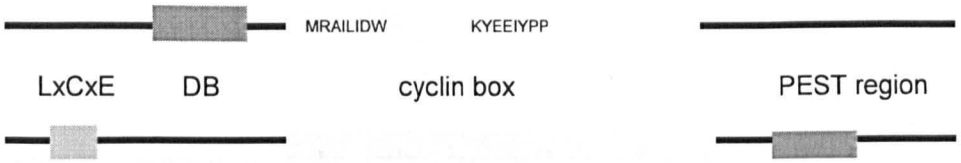


Figure 1.3 Cyclin conserved regions

The cyclin box, destruction box (DB), and the PEST-rich region are drawn as rectangles. The cyclin box is a highly conserved region, found in all cyclins and is believed to be critical for the interaction between cyclins and their kinase partners. The M(R/S)AI(L/F)(V/I/M)DW and KYEE(I/M)Y(A/P/S/T)P sequences are found in almost all mitotic cyclins. The destruction box has a RXALG(N/D/E)IXN motif which triggers the proteolysis of the cyclins during anaphase (A and B-type cyclins). The PEST-rich sequences are found in cyclins C, E and D-type and are thought to render the proteins highly unstable (after Peters, 1994). The N terminal LxCxE motif of cyclins D is involved in Rb binding.

The G1 cyclins do not have destruction box but possess PEST–rich regions that serve as proteolytic signals upon phosphorylation (Rechsteiner and Rogers, 1996). A simplified diagram of the two structural cyclin groups is presented in figure 1.3. Apart from catalysing the activity of the CDK subunits, the cyclins also target the CDKs to specific substrates or subcellular localisations. For instance, cyclin A is localised in the nucleus and cyclins B1 and B2 are located in the cytoplasm because they have a cytoplasmic retention signal (CRS) at their C-terminal ends (Pines and Hunter, 1994). The localisation to the nucleus of CDK1/CyclinB complexes is mediated via phosphorylation within the CRS. Recently it was suggested that the export to the nucleus of CDK1/CyclinB1 could be mediated by the NLS (nuclear localisation signal) of cyclin F which interacts directly with cyclin B1 (Kong et al., 2000). The D-type cyclins which act as transmitters of mitogenic signals to the cell cycle apparatus contain the LxCxE motif that makes the CDK/CyclinD complexes specific Rb kinases (reviewed by Pines and Hunter, 1994). The Rb regulatory pathway is discussed further on in the text.

1.2.2.3 Regulation of the CDK activity by phosphorylation

The CDKs contain several consensus phosphorylation sites that have been shown to be crucial for the activity of the CDK/Cyclin complexes (Russo et al., 1996). An activatory phosphorylation at a conserved Thr residue in the catalytic T-loop (Thr161 in CDK1, Thr 169 in CDK2, Thr 172 in CDK4) by a CAK (CDK-Activating Kinase, itself a complex of CDK7/CyclinH) enables CDK/cyclin complexes to acquire full activity and to perform their cell cycle specific functions (DeBondt et al., 1993; Kato et al., 1993; Gould et al., 1991).

Mutation of the phosphoacceptor Thr abolishes kinase activity of the CDK (Gould et al., 1991). For example, the phosphorylation of the CDK2/Cyclin A complexes causes a change in the position of the T-loop and affects the putative substrate binding site (Russo et al., 1996).

Inhibitory phosphorylation at the amino terminal site of the kinase domain (Thr14 and Tyr15 in CDK1 and CDK2) is very important in the control of CDK1 activation at mitosis. Fully assembled CDK/Cyclin B complexes will be inactive until the abrupt dephosphorylation of the Thr14-Tyr15 residues at the G2/M boundary. The kinase that phosphorylates Thr15 *in vivo* is the product of the *wee1* gene, originally identified from *S. pombe* (Featherstone and Russell, 1991). A dual specificity phosphatase, Cdc25 dephosphorylates Thr14-Tyr15 bringing the CDK to its fully active state (Berry and Gould, 1996). The regulators of CDK phosphorylation themselves are under controls similar to these described for the CDKs.

1.2.2.4 CDK inhibitors

The CDK inhibitors (CKIs) are normally small (14-27 kDa) proteins that bind and inactivate different CDK/Cyclin complexes in response to a variety of antireplicative stimuli, for instance senescence ($p16^{INK4a}$) or terminal differentiation ($p18^{INK4c}$ and $p19^{INK4d}$). There are two families of CKIs, the INK4 family (Inhibitors of CDK4) ($p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4c}$, $p19^{INK4d}$) and the KIP family ($p21^{CIP/WAF1}$, $p27^{KIP1}$, and $p57^{KIP2}$). The INK4 family members are specific CDK4/6 inhibitors, whereas the KIP members affect the activities of cyclin D-, E- and A-dependent kinases (Sherr and Roberts, 1999).

In normal cells, $p21^{CIP/WAF1}$ is part of the active cyclin-CDK complex. Inactivation of the CDK is achieved by increasing the ratio of bound p21 subunits (Zeng et al., 1997). The expression of p21/p27 is regulated by external stimuli responsible for differentiation-associated growth arrest in the G1 phase.

The INK4 proteins, on the other hand, form binary complexes with CDK4/CDK6 and compete for binding with the activatory cyclin D subunits (Ragione et al., 1996). The crystal structure of several CDK-CyclinD–INK complexes have been resolved and they have confirmed the view that INK4 binding weakens the cyclin's affinity for the CDK by distorting the ATP binding site and misaligning the catalytic residues (Jeffrey et al., 2000).

1.3 Plant CDKs and regulators of the plant cell cycle

The recognition of the importance of a CDK in the control of the G1 to S and G2 to M transitions in yeast cells (*Cdc2* in *Sch. pombe* and CDC28 in *S. cerevisiae*), and the discovery that a related protein is a major component of the MPF (mitosis promoting factor) of metazoan eggs and embryos (Nurse, 1990), prompted active research on the identification and characterisation of CDKs from a variety of organisms, including plants. A number of plant cdc2-like genes were identified originally utilising PCR and primers designed to conserved regions of the CDK catalytic domain. The same approach was used for the discovery of many plant cyclin-like genes. Later on, functional complementation of yeast G1 cyclin mutants led to the discovery of plant D-type cyclins (Soni et al., 1995; Dahl et al., 1995). At the present stage, we know that plants have extensive families of *CDK* and cyclin genes.

1.3.1 Overview of plant CDKs

At primary structural level, the plant CDKs are showing the conserved characteristics of the eukaryotic CDK family. The latest classification of the plant CDK-related kinases orders them into 5 groups, *CDKA* to *CDKE* according to certain structural characteristics, most importantly, the type of PSTAIRE motif in the α1 helix (Joubès et al., 2000). Thus, genes with the PSTAIRE motif are considered as one family – *CDKA*. In terms of the *CDKA*s' functional characteristics, they have been shown to complement *cdc2/cdc28*^{ts} mutants in yeast (see table 1.2). At the transcriptional and protein level these proteins are quite stable with transcripts being detected in non-dividing tissues and cells (Martinez et al., 1992; Hemerly et al., 1993). However, the kinase activity of these enzymes peaks during G1/S and then at G2/M transitions (Reicheld et al., 1999; Sorrel et al., 2001). Therefore, it has been proposed that this class of CDKs are not only involved directly in the processes of cell division, but also maintain the competence of quiescent and differentiated cells for division (Hemerly et al., 1993; Mironov et al., 1999). The *CDK-A* type associate with different cyclin partners (cyclins A, B and D- type) to perform specific functions at G1/S and G2/M transitions. For example, a RBR (pRb related protein) phosphorylating activity can be isolated from different plants including *Arabidopsis*, maize, wheat and tobacco. In *Arabidopsis* this activity consists of the PSTAIRE CDK A/cyclin D2 complex. Not surprisingly, this activity is at its highest during G1 and early S phases and declines progressively towards G2 in synchronised tobacco BY2 cell suspensions (Boniotti et al., 2001). Table 1.2 recapitulates on some of the experimental data about plant CDKs.

The *CDKB* family is defined by the PPTALRE (*CDKB1*) or PPTTLRE (*CDKB2*) motif which is plant specific (Joubès et al., 2000). Representatives from this family could not bind yeast temperature sensitive *cdc2/cdc28* mutants (Imajuku et al., 1992; Fobert et al., 1996). The expression pattern of the two alfalfa *CDKB*-type kinases as well as several phosphorylation experiments demonstrate increased activity during G2 and M phase of the cell cycle (Magyar et al., 1997; Mészáros et al., 2000). A *CDKB* type activity was described in tobacco BY2 cell suspension (Porceddu et al., 2001) and it was demonstrated that the protein peaks at G2/M transition followed by a peak of the corresponding PPTALRE CDK-associated histone H1 kinase activity in the middle of M phase. It has been suggested that Y15 phosphorylation event may control the timing of the kinase activity, and it was further demonstrated that in monomeric form this tobacco B-type CDK has only negligible activity.

The genes forming the two sub-groups of *CDKBs* have slightly different patterns of expression. *CDKB1* transcripts accumulate in all phases of the cell cycle with the exception of G1, whereas *CDKB2* expression is specific to the G2 and M phases (Joubès et al., 2000). The tomato Lyesc;*CDKB1;1* and *CDKB2;1* can be detected in young fruit up to 15 days after pollination when endoreduplication occurs. The disappearance of the signal also supports the view that the B-type CDKs have a role in the G2/M phase (Joubès et al., 2001). In general, it is accepted that the expression pattern of this group of CDKs is dependent on the cell cycle stage, in contrast to the *CDKAs* (reviewed in Mironov et al., 1999).

The remaining three families of plant cyclin-dependent kinases, *CDKC*, *CDKD*, and *CDKE* each have only a few representatives to date. The *CDKC* group is characterised by the PITALRE motif, which makes them similar to the human CDK9 (de Falco and Giordano, 1998). There is only limited data about the expression pattern of the *CDKC* group; for instance the Medsa;*CDKC;1* is expressed constitutively in synchronised cell suspensions (Magyar et al., 1997) and Lyces; *CDKC;1* is detected preferentially in dividing tissues in a uniformly distributed signal (Joubès et al. ,*2001*). The involvement of this group of kinases in the regulation of the cell cycle remains to be demonstrated, but it has been suggested that they may not be directly involved in the control of the cell cycle (reviewed in Mironov et al., 1999).

The *CDKD* class has been defined by the presence of the conserved N(I/F)TALRE motif, and is close to the CDK-activating kinase (CAK), CDK7/MO15 (Fesquet et al., 1993). The rice *CDKD;1* has been demonstrated to phosphorylate rice *CDKA;1,* human CDK2, the carboxyl-terminal domain (CTD) of RNA polymerase II of *Arabidopsis* as well as to complement a CAK-deficient mutation of budding yeast (Yamaguchi et al., 1998). The preferential

expression of *CDKD;1* during G1 and S phase in partially synchronised suspension cells produce further proof of the involvement of this kinase in cell cycle regulation (Sauter, 1997). The work of Masatoshi et al., 2000 demonstrated a specific interaction between rice CAK and a H-type cyclin (*Os;CycH;1*) in yeast cells, as well as *in vitro* binding of the kinase to *Os;CycH;1* and *Os;CycC;1*. Interestingly Orysa;*CDKD;1* did not bind to A- and B-type rice cyclins. Moreover, cyclin H positively regulates the phosphorylation activity of the rice CAK on both CTD and CDKs. Thus it seems that *Os;CycH;1* is a specific partner of the rice CAK.

The last defined group of CDKs in plants, the *CDKE* type contains a single sequence from alfalfa, Medsa;*CDKE;1* (Magyar et al., 1997), and is characterised by the presence of an SPTAIRE motif in the α1 helix. During cell cycle progression in a synchronised cell suspension it produced a weak constitutive signal, and it is yet to be demonstrated that it has an involvement in cell cycle regulation.

The current classification of the plant CDKs may need to be readjusted in the future as more data about the partners of these kinases, and their activities is obtained. On a strictly structural basis, the *CDKD* and *CDKE* families may also need to be redefined, as genes from more plant species become available.

However, some additional plant CDK-like genes have not been included in the latest classification because they do not posses enough structural similarities to other defined CDKs or because of the limited experimental data that exist at the moment about their properties. For instance, the cak1At gene from *Arabidopsis*, which possesses CDK-phosphorylation activity but not activity towards the CTD of RNA polymerase II, was isolated as a suppresser of temperature-sensitive *cak* mutant of budding yeast (Table 1.2; Umeda et al., 1998). Another interesting CDK-like gene is the *Arabidopsis* cdc2cAt (Lessard et al., 1999). It belongs to a subfamily of at least 10 CDK-like genes, which appear to be specific for plants. Interestingly these sequences display typical cyclin-dependent kinase overall primary structure but have a more convergent 'PSTAIRE' epitope of the type ESV(K/R)FMARE. There is no further data about the role of these genes and their connection with the cell cycle machinery, but it is highly likely that at least some of them could exhibit cell cycle related functions. This suggestion is based on observations of the expression pattern of the *Bvcrk1* gene, which is the subject of study in the following chapters of this thesis.

Table 1.2 Summary of the experimental data about some plant CDKs.

gene and accession number	expression	localisation	interactions	complementation of yeast mutants	remarks
			CDK A class		
Arath;CDKA;1 M59198	G1/S/G2/M	pre-prophase band, phragmoplast[1]	Arath;CycD1;1[2] Arath;CycD4;1[2] Arath;CycB1;1[1] Arath;CycB2;1[1] Arath;CycB2;2[1] ICK1At[3] ICK3At[3] ICK2[16]	cdc2[ts 4] G1/S and cdc2/cdc28[ts] G2/M mutants[13]	cell division and re-entry into the cell cycle in complex with cyclin D2 has RBR phosphorylating activity[23]
Antma;CDKA;1 X97637	G1/S/G2/M[15]			Partially complement cdc2-33[ts] mutation[15]	
Antma;CDKA;2 X97638	G1/S/G2/M[15]			Partially complement cdc2-33[ts] mutation[15]	
Glyma;CDKA;1 M93140	Higher transcription in shoot-derived organs			cdc28[ts] mutant[14]	
Glyma;CDKA;2 M93139	Higher transcription in shoot-derived organs			Rescues cdc28[ts] mutant[14] better than CDKA;1	
Lyces; CDKA;1 Y17225			Lyces;CycA2;1[23] Lyces;CycD3;1[23] Lyces;CycB2;1[23]		
Lyces; CDKA;2 Y17226			Lyces;CycA2;1[23] Lyces;CycD3;1[23]		
Medsa;CDKA;1 M58365	S and G2/M[9]		Medtr;CycD5[10] Medtr;CycD4[10] CKS Zm[10] CKS Ms[10]	cdc28[ts] G2/M[5]	
Medsa;CDKA;2 X70707	S and G2/M[9]		Medtr;CycD5[10]	cdc28[ts] G1/S[5]	
Nicta;CDKA;3 D50738	G1/S/G2/M[17]		Nicta;CycD3;1[6] Nicta;CycA;1[6] (Ntcyc25)	cdc28[ts] G1/S[17] cdc28[ts] G2/M[17]	Cdc2/cycD phosphorylates NtRb[6] in vitro
Nicta;CDKA;1 AF28947					
Orysa;CDKA;1 X60374	G1/S/G2/M[18]			cdc28[ts] G2/M[11]	Phosphorylated by rice CAK on T161 in vitro[19]
Orysa;CDKA;2 X60375	G1 and S[18]			fails to rescue cdc28[ts] G2/M[11]	Not phosphorylated by rice CAK in vitro[19]
Zeama;CDKA;1 A40444				rescues G2/M cdc28[ts 12]	
			CDK B class		
Arath;CDKB1;1 D10851	S and G2/M[7]		Arath;CycB1;1[1] Arath;CycB2;2[1] CKS1aAt[2]	Fails to complement cdc2/cdc28[ts]	
Antma;CDKB1;1 X97639 Nicta;CDKB1;1	S and G2/M[15]			Do not complement cdc2-33[ts] mutation[15]	
Antma;CDKB2;1 X97640	G2/M[15]			Do not complement cdc2-33[ts] mutation[15]	
Arath;CDKB2;2	G2/M[24]	expresion limited to cells in mitosis[24]			
Lyces; CDKB1;1 AJ297916					
Lyces; CDKB2;1 AJ297917			Lyces;CycA2;1[23]		
Medsa;CDKB1;1 X97315	S and G2/M[9]		CKS Zm[10] Medsa;CycD4[10]		
Medsa;CDKB2;1 X97317	G2/M[10]	pre-prophase band, spindle, metaphase spindle & phragmoplast[10]	Myosin Ms[10] CKS Zm[10] Medsa;CycD4[10]		
Orysa;CDKB;1 D64036	G2/M[21]				Not phosphorylated by rice CAK in vitro[19]

Table 1.2 continued

gene and accession number	expression	localisation	interactions	complementation of yeast mutants	remarks
			CDKC, CDKD and CDKE class		
Arath;CDKC;1 T42526					
Arath;CDKC;2 T20748					
Lyces;CDKC;1 AJ294903	floral meristems, young fruit				suggested non-cell cycle regulation
Medsa;CDKC;1 X97314					
Orysa;CDKD;1 X58194	G1/S[18]		Os;cycH1;[20] Pt;cycH;1	fails to rescue cdc28[ts] G2/M[11] overcomes civ1-4 mutation and suppresses civ1/cak1[ts] mutation in S. cerevisiae[19]	CAK activity[18] Phosphorylates CDK2, CTD of At RNA Pol II[19], function enhanced when in complex with Os;cycH;1[20]
Medsa;CDKE;1 X97316					
			CDK-related sequences		
CAK2At AB009399				S. cerevisiae cak1/civ1[ts] mcs6 (crk1/mop1)	phosphorylates T160 of CDK2 but does not phosphorylate CTD of the largest subunit of RNA polymerase II
Cdc2At		restricted to flowers[22]			

Sources of reference for table 1.2:

[1] Stals et al., 2000; [2] De Veylder et al., 1999; [3] De Veylder et al., 1997; [4] Ferreira et al., 1991; [5] Hirt et al., 1993; [6] Nakagami et al., 1999; [7] Segers et al., 1996; [8] De Veylder et al., 1997; [9] Magyar et al., 1997; [10] Mészáros et al., 2000; [11] Hashimoto et al., 1992; [12] Colasanti et al., 1991; [13] Imajuku et al., 1992; [14] Miao et al., 1993; [15] Fobert et al., 1996; [16] Lui et al., 2000; [17] Setiady et al., 1996; [18] Sauter, 1997; [19] Yamaguchi et al., 1998; [20] Yamaguchi et al., 2000; [21] Umeda et al., 1999; [22]Lessard et al., 1999; [23] Joubès et al., 2001; [23] Boniotti et al., 2001; [24]Menges and Murray (2002)

1.4 The G1/S and G2/M phase transitions and regulators of the plant cell cycle

1.4.1 The G1/S controls

The replication of the DNA requires that certain conditions be in place for the process to start. Most importantly, the enzymes of the replication machinery have to be activated, and the origins of initiation of DNA replication must be in place.

In yeast, one single protein kinase (CDC2/CDC28) binds consecutively to G1 and S phase-specific cyclins, thus driving the G1/S progression (Nasmyth, 1993). In mammalian cells, several G1 and S phase-specific cyclins activate various CDKs at the G1/S transition. Activation of cyclin E-CDK2 at the late G1 restriction point and activation of cyclin A-CDK2 at the G1/S transition suggest the involvement of these cyclin-CDK complexes at specific cell cycle regulatory checkpoints (Elledge, 1996). In contrast, the D-type cyclins are inactive in

G0 but become activated by growth factors early in G1 (Matsushime et al., 1994) before the activation of the cyclin E and cyclin A associated kinase activities.

One important negative regulator of the replication apparatus is the pRb protein. Originally identified as retinoblastoma tumour suppressor protein, the pRb belongs to a small family of so called pocket proteins (pRb, p130, and p107) which share two structurally similar domains A and B that form the pocket. The pRb can interact via the pocket with variety of cellular proteins and has role not only in the negative regulation of the G1/S transition but also in processes of tissue differentiation (Adams, 2001). pRb negatively regulates the G1/S phase progression by binding to transcriptional factors such as E2F family members, and histone remodelling proteins like the histone deacetylases (HDAC) (de Gregori et al., 1997). E2F regulates transcription of genes involved in DNA synthesis and cell cycle regulation. Moreover, a study that combined immunoprecipitation of cross-linked protein-DNA complexes with DNA microarray analysis has identified many E2F target genes which not only have role in the DNA replication and cell cycle (Rb, Cdk1, Cdc25A, cyclin A2, E2F2-3, Cdk2, CDC6, DNApolα) but also genes involved in cell cycle checkpoint control (MAD2, p53), DNA repair (RAD54, PCNA), and chromatin modification (members of the SMC and Histone gene families) (Ren et al., 2002). This data indicates that E2F has a broader role in the cell cycle regulation. When bound to Rb, E2F acts as a repressor and the mechanism of pRb suppression is dependent on the phosphorylation state of the protein. pRb contains 16 potential phosphorylation sites for CDKs. The early phosphorylation of the pRb is done by cyclin D-CDK4/6 (Ezhevsky et al., 2001) which allows the pRb to bind to the E2F transcription factors so that the E2F responsive genes remain transcriptionally inactive in early G1 phase. Later in G1, pRb is phosphorylated by cyclin E/CDK2 thus releasing E2F to drive the transcription of its responsive genes (figure 1.4). The pRb phosphorylation is maintained until entry into M phase, where it is dephosphorylated by the protein phosphatase PP1a and is maintained in underphosphorylated state until cells are stimulated to replicate again (Bollen and Belleus, 2002). Primary target of the pRb in G1 is the E2F transcription factor, a heterodimeric complex containing the E2F and the DP (dimerisation partner) subunits. The two subunits contain conserved DNA binding and dimerisation domains. After dimerisation, the E2F/DP can bind DNA with high sequence specificity. The E2Fs typically contain a strong C-terminal activation domain, which is masked in the E2F/DP/Rb complexes hence, their repressive function.

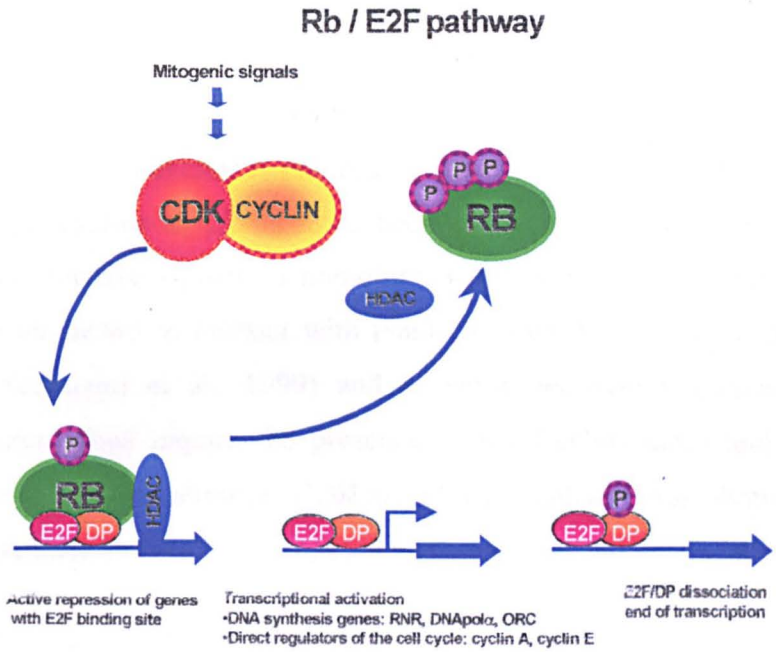


Figure 1.4 The Rb/E2F pathway

In early G1 phase Rb is bound to the E2F/DP heterodimer and HDAC thus preventing transcription of E2F responsive genes. The Rb interacts with E2F/DP via the transactivation domain of E2F. Usually HDAC interacts with Rb via a LxCxE motif. Upon mitogenic stimulation an active CDK/cyclin complex forms that hyperphosphorylates Rb thus dissociating HDAC and Rb from the E2F/DP transcription sites. Active transcription from the E2F responsive sites continues until the phosphorylation of the DP subunit in late G1.

The sequential phosphorylation of pRb by G1/S cyclin/CDK complexes reverses the pRb repression and drives the cells into S-phase entry.

Similar control mechanisms operate in the plant G1 to S phase transition. The progression through the cell cycle is controlled by specific CDKs in complex with different regulatory cyclin subunits. At the G1/S phase transition two principal control mechanisms can be described. On one hand, early in G1, the D-type cyclins act as mediators and integrators of environmental and developmental cues to the cell cycle machinery (reviewed in Shen, 2000; Oakenfull et al., 2002). On the other, the Rb/E2F regulatory pathway is seen to negatively regulate S phase-entry.

1.4.1.1 Cyclin D as link between internal and external signals

Originally, the first cyclin Ds form plants were isolated by yeast complementation (Soni et al., 1995), and today 10 members are found in the *Arabidopsis* genome. Currently it is proposed that they are separated into at least six separate cyclin D groups, which may also have different functions (Oakenfull et al., 2002). Structurally, the D-type cyclins can be defined by the presence of the LxCxE motif at the N terminal part of the protein. This motif is believed to take part in the binding to the Rb protein (reviewed by Pines and Hunter, 1994; Oakenfull et al., 2002). In this way the cyclins target their kinase partners to the Rb for phosphorylation and thus represent a link between external signals and the core cell cycle regulatory unit. It

has been shown that in plants cyclin Ds levels and activity are sensitive to the presence of hormones and sugars (reviewed in Oakenfull et al., 2002). For instance, the expression of cyclin D3;1 is induced upon treatment with cytokinins before the G1/S transition (Soni et al., 1995), and the expression of cyclin D3;1 in transgenic *Arabidopsis* results in cytokinin-independent callus growth (Riou-Khamlichi et al., 1999). The expression of two other *Arabidopsis* cyclins is induced by sucrose (Soni et al., 1995; De Veylder et al., 1999). Thus D type cyclins in plants have been suggested to act as mediators of internal and external proliferative signals in homology with their animal counterparts. The D-type cyclins have been shown to interact with plant Rb proteins in *in vitro* protein-protein interaction assays (Nakagami et al., 1999) and in yeast two-hybrid system (Huntley et al., 1998). These interactions require the presence of the LxCxE motif and the pocket domains of the Rb proteins. Furthermore, CDKA/cyclin D2 activity was shown to be a RBR phosphorylating complex in vivo (Boniotti et al., 2001).

1.4.1.2 The Rb/E2F

In the last few years it was demonstrated that plants, like other multicellular eukaryots posses homologues of the Rb and E2F proteins. Homologues of the pRb proteins have been identified from maize (Grafi et al., 1996; Ach et al. 1997), *Nicotiana tabacum* (Nakagami et al., 1999), *Arabidopsis thaliana* (Kong et al., 2000). It appears that maize has up to three family members, while a single Rb gene has been defined in *Arabidopsis* (Durfee et al., 2000; Vandepoele et al., 2002). The Rb have been shown to interact with plant viral proteins (Grafi et al., 1996; Xie et al., 1996). Plant D-type cyclins bind the maize RRB1 via the LxCxE motif (Arch et al., 1997; Huntley et al., 1998). Plant Rb has been shown to bind to E2F proteins from wheat, and tobacco (NtE2F), (Ramirez-Para et al., 1999; Sekine et al., 1999). The plant Rb is also regulated by phosphorylation with cyclin D/CDK complexes being the major kinase (Nakagami et al., 1999). Finally, it has been shown that plant Rb interacts with E2F proteins from wheat, *Arabidopsis*, tobacco and human (Huntley et al., 1998; Ramirez-Para et al., 1999; Sekine et al., 1999) via the transactivation domain of the E2F transcription factor.

As it was mentioned earlier in the text, the Rb function in the cell cycle regulation is to suppress the transcription of E2F responsive genes. By binding to the trans-activation domain of the E2F, Rb blocks transcription of the E2F responsive genes. Indeed, E2F binding sites were found in the promoters of the tobacco RNR gene, the *Arabidopsis cdc6* and MCM genes (Chaboute et al., 2000, de Jager et al., 2001; Stevens et al., 2002). Potential E2F binding sites

exist in many other genes which are thought to take part in the G1/S cell cycle progression (de Jaeger et al, 2001).

E2F genes have been isolated from several plant species, including tobacco, wheat and *Arabidopsis*. Genome-wide analysis of the core cell cycle genes in *Arabidopsis* has confirmed that there are 8 members of the E2F/DP family. Three genes are most closely related the animal E2Fs (E2F1/E2Fb; E2F2/E2Fc; E2F3/E2Fa), two belong to the DP family group (DPa and DPb), and further three genes belong to a newly defined group of DP-E2F related genes or DELs, which have been also called ELP (E2F/DP like protein), or E2L (E2F like). Structurally, the E2Fs and DPs share homology over their DNA binding and dimerisation domains, and the E2Fs also share the marked box and RB-binding and transactivation domains, similarly to their animal counterparts. In contrast, the DELs contain two separate domains with limited homology to the DNA binding domains of E2F and DP. It was suggested that they could bind to DNA as monomers and serve as transcriptional repressors by competing for DNA binding sites with the E2F/DP transcription factor (Kosugi and Ohashi, 2002). E2F and DP proteins have been shown to heterodimerize in vitro (Magyar et al., 2000) and to interact in the yeast two hybrid system (Magyar et al., 2000; de Jager et al., 2001). High affinity binding to a DNA probe occurs when AtE2F and AtDPs are co-translated *in vitro*, indicating that the DNA binding activity depends on the formation of a dimer (Kosugi and Ohashi, 2002). Furthermore, the AtE2F1 can bind the promoter of the S-phase-regulated gene *cdc6* (de Jager et al., 2001) and E2F elements have been implicated in the up-regulation of the genes for ribonucleotide reductase RNR-2 and PCNA (Chaboute et al., 2000; Egelkrout et al., 2001). Overexpression of E2Fa/DPa in *Arabidopsis* results in uncontrolled cell proliferation and delayed differentiation. Taken together, it can be concluded that the E2F/DP play a role in the control of S-phase by up-regulating genes necessary for cell cycle progression and Rb and DELs independently suppress the onset of S phase.

1.4.1.3 The role of CDK inhibitors

Homologues of the p21/p27 type of inhibitors have been described for plants but no homologues of the INK family have been reported. *Arabidopsis* contains 7 genes but information about their function is still limited. For instance, ICK1 can bind CDKA and cyclin D3 in yeast two-hybrid interactions, and it also can inhibit kinase activity of CDK *in vitro* (Wang et al., 1998). The senescence hormone ABA was found to induce the transcription of ICK1 which leads to a reduction in cell division with a pronounced dwarfing

effect (Wang et al., 2000). ICK2/KRP overexpression results in slower cell division but without disruption of the overall body plan (De Veylder et al., 2001).

An insight into the timing of KRP (Kip-related proteins) expression comes from a study of synchronised *Arabidopsis* cells and it shows that the KRPs are expressed in different intervals during G1/S (Menges and Murray, 2002). The KRP2 (ICK2) was detected in early G1 phase in cells re-entering the cell cycle after sucrose starvation, while KRP3 transcript appeared in G1 and steadily increased to a maximum in G2. The authors propose a role for KRP3 in the control of the transit between G2 and mitosis while KRP2 is specifically involved in the re-entry of cells into division. A third inhibitor, KRP1 was expressed at high level in starved cells and the expression decreased after sucrose addition then increased again at about G2/M which indicates a possible role in non-dividing cells and in mitosis (Menges and Murray, 2002).

1.4.2 G2/M transition in plant cells

Mitosis is universally conserved in all nuclear organisms, although specific features exist in different organisms. During mitosis the replicated chromosomes are split by the mitotic spindle, at the same time as the rest of the cellular content (the organelles, the membranes and the cytoplasm) is divided between the two daughter cells. Key players in the process of preparation and completion of mitosis are the CDKs activated by mitotic cyclins. In animal cells, the $p34^{cdc2}$ –cyclin B complexes catalyse chromosomal condensation and nuclear envelope breakdown during mitosis, thus having a central and rate-limiting function in the transition from G2 into M phase (Norbury and Nurse, 1992). Activation of the Cdk1 (cdc2) at the G2/M transition is precisely regulated through accumulation of cdc2-associated cyclin B and phosphorylation on three sites in Cdk1 (Nurse, 1997). Phosphorylation at Thr161 residue of Cdk1 by cyclin H-Cdk7 complex is essential for Cdk1 catalytic activity (Harper and Elledge, 1998). The inhibitory phosphorylation Y15 (Cdk1) on Cdc2 is carried out by the Wee1 tyrosine kinase and on Thr14 by Myt1 kinase (Liu et al., 1997). Moreover, activation of Cdk1 is initiated at mitosis by increased activity of cdc25C phosphatase that dephosphorylates Cdk1 on Thr14 and Tyr15 (Gautier et al., 1991). Additionally, the Cdk inhibitor $p21^{Cip1}$ was shown to directly inhibit Cdk1 activity (Dulic et al., 1998; Niculescu et al., 1998; Yu et al., 1998). Exit from mitosis begins with the degradation of the cyclin B partner of the CDK.

In the last few years, plant cell cycle research has produced sufficient data about the specific features of plant cell cycle regulation. One obvious difference is that in contrast to animals, plants possess two classes of CDKs active in mitosis. One is CDKA, which seems to have a

role in G1/S as well as in G2/M, if judged by the protein accumulation and kinase activity (reviewed in Joubés et al., 2000). The B-type CDKs are apparently active from the onset of S phase onwards with peak in G2. The mRNA of CDKB1;1 is detectable earlier than that of CDKB2;2 (G2/M phase peak) in *Arabidopsis* cell suspension (Menges and Murray, 2002), similarly to previously reported expression for the alfalfa B-type CDKs (Magyar et al., 1997) (see table 1.2). The proteins of the B-type CDKs also show cell cycle-dependent accumulation pattern. The histone H1-phosphorylating activity of the CDK B is equally G2/M phase specific (reviewed in Stals and Inzé, 2001).

The catalytic partners of the plant CDKs during mitosis are not yet well established although it is logical that they will be those cyclins expressed in the same timeframe. Most plant A and B-type cyclins have G2/M phase specific expression patterns (reviewed in Chaubert-Gigot, 2000; Ito, 2000). The kinase activity associated with cyclin Bs correlates with their protein levels and is highest during G2/M. The cyclin A2-associated kinase activity from *Medicago* was found to be biphasic with a first peak occurring in the S phase and a second major peak during G2/M transition (Roudier et al., 2000). The activation of the CDK is the crucial biochemical step required for mitosis, and is achieved by the timely expression of the cyclin partners. Recently it was shown that the transcription of some of the G2/M phase specific genes is mediated via a cis-acting element, called M-specific activator (MSA). Most notably, The MSA element was found in the promoter of the tobacco cyclin B1 gene, Nicsy;CycB1;1. It was shown that this element is recognised by different Myb-like factors two of which act as activators (NtMybA1 and NtMybA2) and one is a repressor (NtMybB) (Ito et al., 2001). The authors of this work suggest that the Myb-mediated transcriptional regulation can be a feature of many cell cycle genes with a function in the G2/M.

The proteolytic degradation of the cyclin B-type is mediated by the destruction box and is believed to be required for exit from mitosis. A very clear example of the involvement of the D-box pathway in plants comes from experiments with tobacco mitotic cyclins. N-terminal domain fusions of tobacco mitotic cyclins and a reporter protein produce constitutively expressed fusion protein that accumulates in S phase until mitosis when they undergo D-box dependent destruction (Genschik et al, 1998). The timing of the cyclin degradation was observed by time laps images of BY2 cells expressing tobacco cyclin B1-GFP fusions and it starts at the onset of anaphase. A non-degradable cyclin protein remains associated to the chromosomes throughout anaphase (Criqui et al, 2001).

The phosphorylation on the conserved theronine residue in the T-loop of the CDK by CAK is a further requirement for the activation of the kinase. Although plant CAK has been shown to

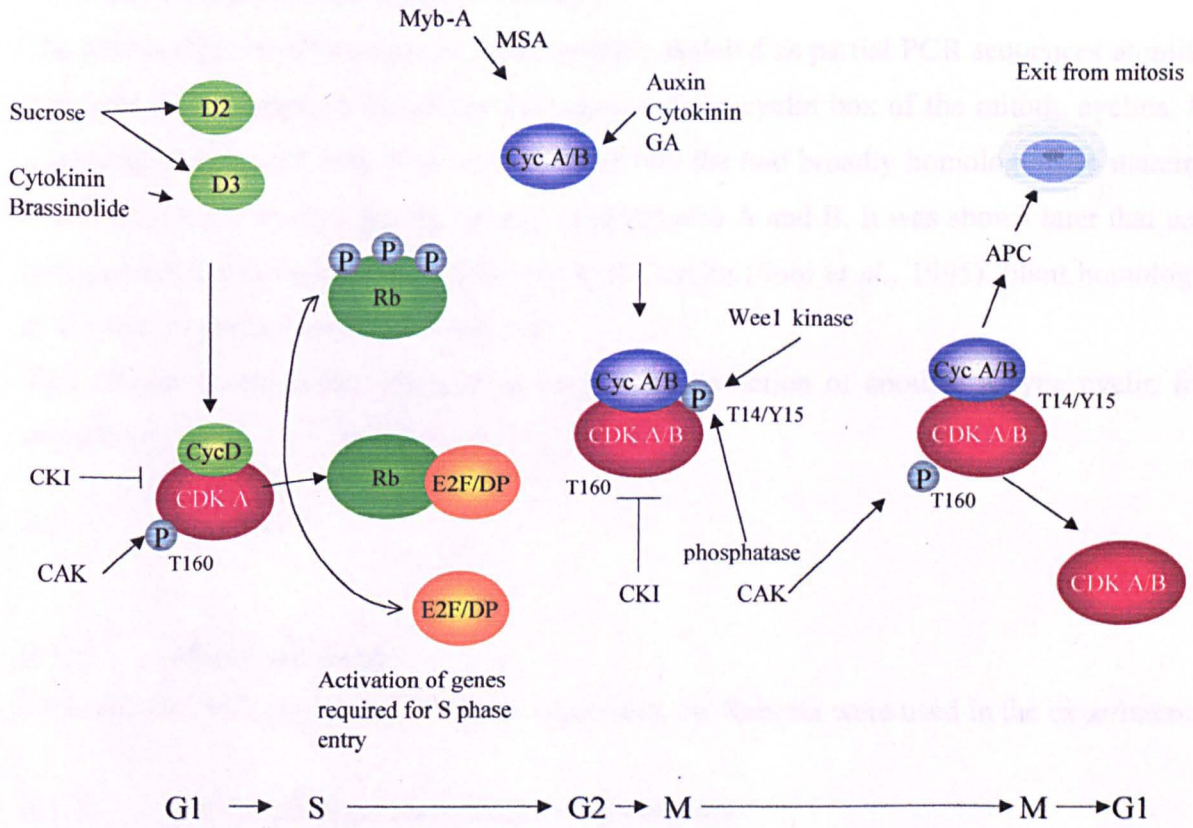
regulate S-phase progression, no mitotic CAK function has been demonstrated so far. Inhibitory phosphorylation by the wee1/myt1 type kinases can be reversed by the activity of cdc25 phosphatase. Plant orthologs of the wee1 gene have been described in maize and in *Arabidopsis* (Sun et al., 1999; Vandepoele et al., 2002) but no cdc25 type phosphatase could be pinpointed. The *Zeama;WEE1* isolated from maize endosperm could inhibit $p13^{Suc1}$ - associated mitotic CDK activity. The transcripts were most abundant in actively dividing tissue, and in the endosperm during the period of endoreduplication (Sun et al., 1999) and could be one of the factors regulating the inhibition of the mitotic CDK activity.

A plant cdc25 functional homologue is very likely to exist even if no homologue was found in the *Arabidopsis* genome (Vandepoele et al., 2002). Recombinant yeast or fruit fly cdc25 can activate tobacco and alfalfa $p13^{Suc1}$ bound CDK *in vitro* (Mészaros et al., 2000; Zhang et al., 1996). In tobacco expressing the yeast cdc25 gene cells divide with a reduced size, an observation that fits to the existence of a phosphatase activity in the onset of mitosis (McKibbin et al., 1998).

In conclusion, a diagram of the cell cycle model in plant is presented in the next page of this chapter.

Figure 1.5: Model for the regulation of the cell cycle in plants

During G1 phase, several growth factors (for instance sucrose, cytokinin) regulate the expression of D- type cyclins. In their turn, the D-type cyclins associate with CDK A to promote the G1/S transaction by triggering the phosphorylation of the Rb. The released E2F transcription factor drives the expression of genes necessary for DNA replication. Later in the cycle, at the G2 the transcriptional activation of cyclins A and B via the MSA (M-specific activator) element. Active CDK/cyclinA/B complexes are released from the inhibitory phosphorylation of *wee1* and further activated by T-loop phosphorylation within the kinase catalytic domain. Targets for the G2/M phase kinase activity are microtubules, G1/S regulatory elements. The activity of the G2/M specific kinases declines following the targeted destruction of their cyclin partners.



CHAPTER 2 PCR BASED ISOLATION OF CYCLIN-LIKE SEQUENCES

The PCR based isolation of a couple of mitotic cyclin partial sequences is presented in this chapter. The results were included in the thesis even though the work on cyclin genes in sugar beet has not been taken very much further than the initial isolation of the partial cyclin box sequences. Chronologically, this is the first piece of experimental work done in the project and it was carried out in close collaboration with a previous PhD student, M. J. Kirby who was at the time finishing her work on sugar beet cell cycle. Later, the cyclin box fragments were used in sugar beet genomic library screening and few positives were identified and left for further characterisation in the laboratory.

The first cyclin genes from plants were typically isolated as partial PCR sequences amplified with primers designed to the conserved regions of the cyclin box of the mitotic cyclins. Not surprisingly, all of the new plant entries fitted into the two broadly homologous to mammals A and B type cyclins groups, designated in plants also A and B. It was shown later that using complementation of yeast strain deficient in G1 cyclin (Soni *et al.*, 1995), plant homologues of the D-type cyclins could be identified.

This chapter presents the initial steps towards the isolation of another A type cyclin from sugar beet.

2.1 Materials

2.1.1 Plant material

Diploid (2n=18) 2 day old seedlings of sugar beet, cv. Roberta were used in the experiments.

2.1.2 Bacterial strains and cloning vectors

E.coli XL1 Blue (Stratagene) were used as hosts of the plasmids pCRScript (Stratagene) and pGEM-T (Promega).

2.1.3 PCR primers

All primers used were synthesised on a PCR MATE DNA Synthesiser (Applied Biosystems). Degenerate oligonucleotide primers corresponding to the cyclin box conserved amino acid sequences MRAI(L/F)(V/I/M)DW, KYEE(I/M)Y(A/P/S/T)P and LE(M/I)LYKA (Hemerly *et al*, 1992.) and referred hereafter as primers A, B and C respectively, were used for PCR

amplification. The sequences of the degenerate primers are shown below. The underlined regions at the 5'-end of the primers represent the introduced restriction sites: Bam H I for primers A and B and Eco R I for primer C. N stands for either A, T, C or G.

DEGENERATE OLIGONUCLEOTIDE PRIMERS[1]	
Primer A	5'-GCAGGATCCATGAG(A/G)GCNAT(C/T)CT(C/T)AT(C/T)GA(C/T)TGG-3'
Primer B	5'-GCAGAATTCAT(C/T)GCNTCNAA(A/G)TA(C/T)GA(A/G)-3'
Primer C	5'-GCAGGATCC(A/G)AG(C/T)TCNA(C/T)NA(A/G)(A/G)TACTTNGC-3'

2.1.4 Reagents, media, solutions, enzymes and kits used

Dynabeads oligo d(T_{25}), Dynal, UK;

All restriction enzymes used were supplied from Promega unless stated otherwise in the text. The composition of used media and solutions is given in Appendix A.

2.2 Methods

2.2.1 Germination of seedlings and collection of samples for mRNA isolation

Seeds were left to germinate in trays layered with moist tissue paper and covered with Clingfilm. The trays were kept in the dark for 2 to 4 days until the seedlings emerged.

Root tips of 1 to 2 mm in length were excised from the seedlings with a razor blade and were immersed at once in liquid nitrogen.

Intact two-day-old seedlings were also collected and stored in liquid nitrogen.

The samples were kept in Eppendorf tubes at –70° C as 100 mg aliquots of tissue ready for mRNA isolation.

2.2.2 mRNA isolation

One hundred micrograms of root tips, cotyledons or intact two-day-old seedlings were used as a starting material for mRNA isolation with Dynabeads Oligo d (T_{25}). Throughout the procedure, the manufacturers instructions for mRNA isolation from crude tissue lysates were followed closely (figure 2.1 presents generalised scheme of the procedure). 1–1.5 μg of mRNA was obtained from each 100 mg tissue per extraction. The Dynabeads were

[1] The sequence of the primers taken from Hemerly *et al.*, 1992

regenerated after each use and reused twice more on the same type of starting tissue homogenate.

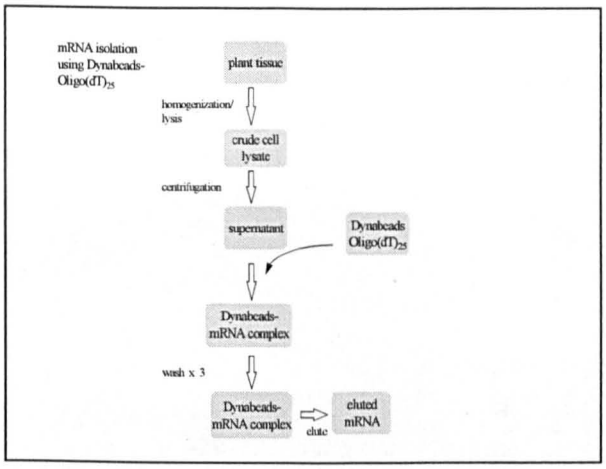


Figure 2.1

Schematic representation of the major steps in mRNA isolation when using Dynabeads oligo d(T_{25}) magnetic beads.

2.2.3 Synthesis of cDNA

Single stranded cDNA was prepared from 1 μg poly (A^+) RNA isolated from root tips and cotyledons. The synthesis of the first strand cDNA was primed with three different primers, namely: oligo ($dT)_{36}$ primer (New England Biolabs), random hexamers (Promega) or with antisense degenerate primer C (designed with homology to the cyclin box).

Reverse transcription reaction		
mRNA		1μg
Reverse transcription primer	Oligo d(T_{36})	80 pmol
	Random hexamers	80 - 100 pmol
	Degenerate template specific primer C	100 pmol
Bring to total volume of 20μl with dH_2O and incubate at 70°C for 10 minutes, cool on ice and add:		
First strand buffer, 5x		4μl (1x)
Dithiothretiol, 0.2M		2μl
$dNTP_s$, 10 mM		1μl (500μM each)
SuperScript RT, $2U\mu l^{-1}$*		1μl (2U)

The reactions were incubated at 37°C for an hour, and then the enzyme was inactivated at 75°C for 10 minutes. Three types of cDNA were obtained from the mRNA depending on the primer used in the reverse transcription reaction (figure 2.2, A).

* SuperScript RT was added to the oligo ($dT)_{36}$ and degenerate primer C primed reactions after the tubes have been incubated at 37°C for 2 minutes. For random hexamer primed reaction, the enzyme was added first, then the tube was incubated at room temperature for 10 minutes.

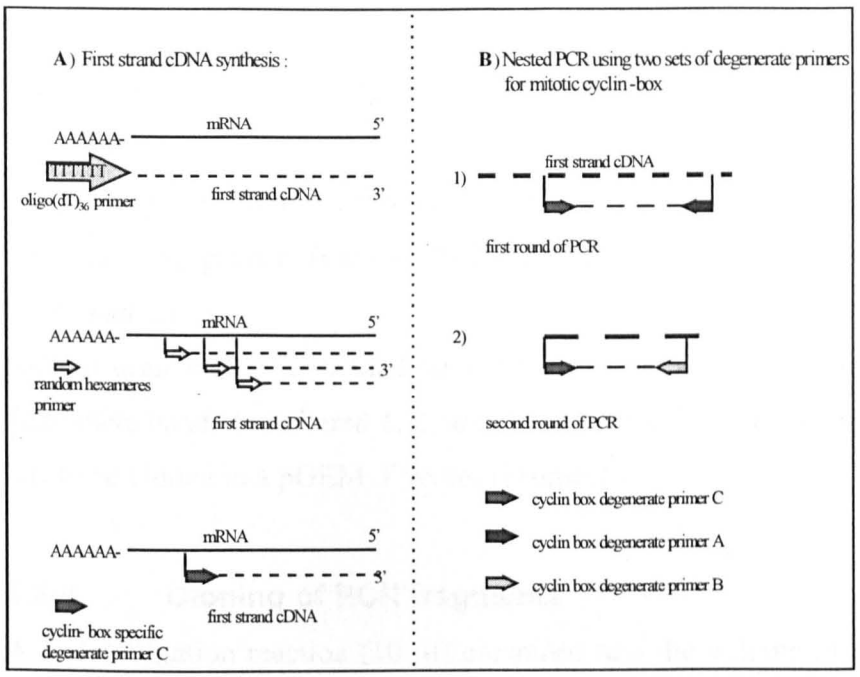


Figure 2.2: cDNA synthesis. (A)-Schematic representation of the cDNA synthesis using different primers for the reverse transcription from mRNA; (B) PCR using degenerate primers with homology to the conserved domains of the "cyclin box".

2.2.4 PCR amplification of sugar beet cyclin -like sequences

Three degenerate oligonucleotide primers based on the conserved regions of mitotic cyclin-box (Hemerly *et al.*, 1992) were used to PCR amplify cyclin-like sequences from sugar beet. First strand cDNA was used as a template for the PCR in a set of two consequent rounds of amplification with the degenerate primers. In the first round of PCR, primers A and C were used, followed by second amplification with primers C and B. The sequence of the degenerate primers is given in the materials section of this chapter.

PCR reaction conditions: The initial denaturation step at 94°C for 1 minute was followed by 44 cycles of amplification at 94°C for 1 minute, 48°C for 2 minutes, and 72°C for 3 minutes. The last amplification had an extension step at 72°C for 10 minutes.

PCR reaction components	
cDNA	5 µl [2]
PCR thermo buffer, 10 x	5 µl (1 x)
$MgCl_2$, 25 mM	4 µl (2 mM)
$dNTP_s$, 10 mM each	1 µl (200 µl)
primer C	1 µl (100 pmol)
primer A	1 µl (100 pmol)
Taq DNA polymerase, $5U.\mu l^{-1}$	0.5 µl (2.5 U)
dH_2O to final volume	50 µl

[2] cDNA synthesized using three different ways of priming - oligo $d(T)_{36}$, random hexamers and gene specific primer C, was used in the PCR amplification. 10^{-2} and 10^{-3} dilutions of the original cDNA were used in the PCR reactions. The same amplification products were obtained with both dilutions.

PCR products were purified from the unincorporated nucleotides, primers and from the polymerase by passing the samples through CHROMA SPIN™-200 H_2O - DEPC columns (Clontech).

A 5 µl aliquot from the purified PCR reaction was taken for a second round of amplification this time using primers B and C. PCR reaction conditions and components were as for the first PCR reaction.

Half of each PCR was loaded on a 2 % 1x TAE agarose gel to separate the products. All distinctive bands (numbered 1, 2, and 3 in figure 2.3) were cut out and purified (Geneclean *II* kit) to be cloned in a pGEM-T vector (Promega).

2.2.5 Cloning of PCR fragments

A typical ligation reaction (10 µl) contained half the volume (4 µl) of the genecleaned PCR band (concentration 50 - 100 ng), 0.5 µl plasmid (conc. 25 ng) 1x ligation buffer, and 1 µl (2 U) T4 DNA ligase (Promega). The reactions were incubated overnight at 4°C.

5 µl of each ligation was used to transform competent *E. coli* XL1 Blue bacteria (Nishimura *et al.*, 1990). The transformed bacteria were selected on LB plates supplemented with 12.5 $\mu g.ml^{-1}$ tetracycline, 50 $\mu g.ml^{-1}$ ampicillin, 1 mM IPTG, 40 $\mu g.ml^{-1}$ X-Gal. 15 -20 white colonies per transformation were picked up and grown in 3 ml LB cultures containing tetracycline and ampicillin (concentration as above). Plasmids were isolated using the plasmid "easyprep method" (Berghammer and Auer, 1993). At this point 1 ml glycerol stocks were made by mixing 0.5 ml from each bacterial culture with 0.5 ml sterile 40% glycerol. The glycerols were stored at -20°C.

About 200 ng from each plasmid were digested with either Sal I + Nco I or Sac I + Sac II restriction enzymes which cut in the multiple cloning site of the plasmid to check for insertions. The clones containing inserts were propagated in 10ml LB cultures and plasmids were for sequencing with Wizard™Plasmid Minipreps (Promega).

Sequenced samples were run on a 6% polyacrylamide gel at 60 watts and dried gels were exposed overnight to Hyperfilm™ - β max (Amersham) X -ray film.

For sequence analyses and homology database searches the GCG programs on SEQNET, Daresbury, UK were used.

2.3 Results

2.3.1 PCR based isolation of cyclin sequences

The PCR based isolation of cyclin-like sequences from cDNA of root tips and cotyledons yielded a couple of very similar sequences. It has to be noted that of all the template cDNAs used in the reverse transcription reaction, only cDNA synthesised with degenerate gene-specific primer C (figure 2.3) produced the expected results.

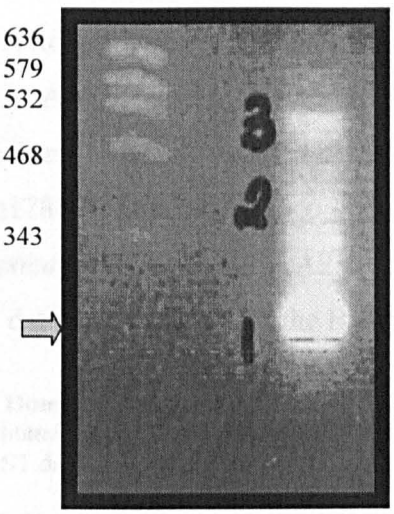


Figure 2.3 PCR amplification of cyclin sequences.
2.0% agarose gel of the PCR products obtained after the second round of amplification using primers B and C. The shaded arrow indicates the PCR product which when sequenced turned out to be a cyclin-like sequence. The size of the DNA markers (bp) is written on the left of the gel picture (marker DNA Lambda digested with Pvu II).

As the PCR products were not subjected to Southern blot analysis, but instead were cloned into plasmids and sequenced, it is not clear whether the other two amplification methods were as unproductive as it seemed. For instance, when random hexamers were used, no amplification products were detected on ethidium bromide stained agarose gel but the oligo-d $(T)_{36}$ primed cDNA gave few bands after the amplification. They were sequenced but showed no homology to the cyclin sequences of interest and were assumed to be non-specific products.

A 257 bp fragment was cloned from root tip cDNA, and 243 bp sequence from cotyledon cDNA. An alignment of the two sequences and a previously isolated cyclin-like fragment (EMBL accession number Z71705) from sugar beet cell suspension (Kirby, 1995) is shown on figure 2.4. It is immediately obvious that the sequences are nearly identical in the amplified region between the conserved regions KYEE(I/M)Y(A/P/S/T)P and LE(M/I)LYKA of the cyclin box. The first cyclin-like sequence in this alignment (Betvu;*CycA*2) has been

previously extended by RACE PCR towards its 3' end (Kirby, 1996). The two other were not, therefore it is difficult to say exactly how similar the three sequences are. However, the few differences observed in the DNAs account for couple of amino acid substitutions in the compared region. This may be an indication of the existence of a multiple group of the same class of cyclins present in the sugar beet genome.

Based on the % homology with other plant cyclins over the partial cyclin core region, the two sugar beet fragments (named here Bv*Cyc*A2(1) and Bv*Cyc*A2(2)), together with the previously described Betvu;*Cyc*A2, fall into the mitotic A2-type class of cyclins as described in Renaudin *at al.*, 1996.

Thus, the closest homologues of the two sugar beet partial sequences are the soybean mitotic cyclin Glyma;*Cyc*A2 (accession number d50869) with 67% identity and 82% similarity, the *Arabidopsis thaliana* Arath;*Cyc*A2;2 (z31402) with 63% identity and Arath;*Cyc*A2;4 with 61% identity (u17889), also the tobacco Nicta;*Cyc*A2;1 (d50869) with 60% identity and 80% similarity, *Brassica napus* Brana; *Cyc*A2;1 (l25405). The table below gives few of the entries from a BLAST database search with the Bv*CycA*2(2) as a query.

Table 2.1 Homology (%) and amino acid alignments of BvcycA2(2) partial sequence to plant A2-type cyclins. The homology encompasses a part of the cyclin core (about 220 aa). The entries in the table are taken from a BLAST database search.

Glycine max D50869 I 67%, P 82%	Query: 2 EEMSAPQVEDFCYITANTYAREEVLNMERKVLNFLCFQLSDPTIKTFLRRYVHAAQATYE 181 EEM AP+VE+FC+IT NTY +EEVL MER+VLN + FQLS PTIKTFLRR++ AAQ++Y+ Sbjct: 295 EEMCAPRVEEFCFITDNTYTKEEVLKMEREVLNLVHFQLSVPTIKTFLRRFIQAAQSSYK 354 Query: 182 DSLVDLEFLAKYLVEL 229 V+LEFLA YL EL Sbjct: 355 APYVELEFLANYLAEL 370	
A.thaliana Z31402 I 63%, P 82%	Query: 2 EEMSAPQVEDFCYITANTYAREEVLNMERKVLNFLCFQLSDPTIKTFLRRYVHAAQATYE 181 EE+SAP VE+FC+ITANTY R EVL+ME ++LNF+ F+LS PT KTFLRR++ AAQA+Y+ Sbjct: 263 EELSAPGVEEFCFITANTYTRREVLSMEIQILNFVHFRLSVPTTKTFLRRFIKAAQASYK 322 Query: 182 DSLVDLEFLAKYLVEL 229 ++LE+LA YL EL Sbjct: 323 VPFIELEYLANYLAEL 338	
A.thaliana I 60%, P 80%	Query: 2 EEMSAPQVEDFCYITANTYAREEVLNMERKVLNFLCFQLSDPTIKTFLRRYVHAAQATYE 181 EE+SAP VE+FC+ITANTY R EVL+ME ++LNF+ F+LS PT TFLRR++ AQA+Y+ Sbjct: 140 EELSAPGVEEFCFITANTYTRPEVLSMEIQILNFVHFRLSVPTTTTFLRRFIQPAQASYK 199 Query: 182 DSLVDLEFLAKYLVEL 229 ++LE+LA YL EL Sbjct: 200 VPFIELEYLANYLAEL 215	
N.tabacum D50736 I 60%, P 80%	Query: 2 EEMSAPQVEDFCYITANTYAREEVLNMERKVLNFLCFQLSDPTIKTFLRRYVHAAQATYE 181 EE+ AP+VE+FC+IT NTY++EEV+ ME +VLN L FQL+ PT K FLRR++ AAQA+Y+ Sbjct: 322 EEICAPRVEEFCFITDNTYSKEEVIKMESRVLNLLSFQLASPTTKKFLRRFIQAAQASYK 381 Query: 182 DSLVDLEFLAKYLVEL 229 V+LEF+A YL EL Sbjct: 382 VPSVELEFMANYLAEL 397	

Further work is needed to show how closely related are the genes corresponding to the three cyclin sequences. The expression patterns of other A-2 type cyclins indicate for a role of these genes in S and G2/M phases of the cell cycle in association with CDK A type. By analogy, it could be assumed that the sugar beet genes would have similar functions in the cell cycle. It

would be very intriguing to demonstrate the link between the cell cycle genes and formation and developmental of sugar beet tap root as this organ has specific features of morphology not observed in the other plant species where cell cycle research is more advanced.

Figure 2.4. A PRETTYBOX alignment of the sugar beet cyclinA-like sequences.
Betvu;CycA2 is a 360 bp sequence isolated from sugar beet cell suspension (Kirby, PhD Thesis, 1996). *Betvu;CycA2(1)*, and *Betvu;CycA2(2)* are the partial cyclin sequences isolated from cotyledons and root cDNA respectively. The alignment was done using the GCG programme package available on SEQNET, Daresbury. UK.

```
                1                                                   50
Betvu;CycA2     AGGTCTCTGA AGAGTACAAA CTGGTCCCTG ACACTCTTTA CCTGACTGTA
Betvu;CycA2(1)  .......... .......... .......... .......... ..........
Betvu;CycA2(2)  .......... .......... .......... .......... ..........

               51                                                  100
Betvu;CycA2     AATCTCATTG ATCGTTTTCT GTCTGGAAAT TACCTGGAAA AACAAAAACT
Betvu;CycA2(1)  .......... .......... .......... .......... ..........
Betvu;CycA2(2)  .......... .......... .......... .......... ..........

               101                                                 150
Betvu;CycA2     GCAGCTCCTG GGAGTAACCT GCATGTTGAT TGCCTCAAAG TACGAAGAAG
Betvu;CycA2(1)  .......... .......... .......... .......... ..CGAAGAAA
Betvu;CycA2(2)  .......... .......... .......... .......... ...gaagaaa

               151                                                 200
Betvu;CycA2     TGAGT..CCC CCAAGTTGAA GACTTTTGCT ACATCACAGC TAACACTTAT
Betvu;CycA2(1)  TGAGTGCCCC CCAAGTTGAA GACTTTTGCT ACATCACAGC TAACACTTAT
Betvu;CycA2(2)  tgagtg.ccc ccaagttgaa gactcttgct acatcacagc taacacttat

               201                                                 250
Betvu;CycA2     GCCAGAGAAG AGGTATTGAA TATGGAGAGG AAAGTTCTCA ACTTTCTCTG
Betvu;CycA2(1)  GCCAGAGAAG AGGTATTGAA TATGGAGAGG AAAGTTCTCA ACTTTCTCTG
Betvu;CycA2(2)  gccagagaag aggtattgaa tatggagagg aaagttctcg gctttctctg

               251                                                 300
Betvu;CycA2     CTTTCAACTA TCTGTACCCA CCATAAAAAC ATTTTTGAGG AGATATGTCC
Betvu;CycA2(1)  CTTTCAACTA TCTGATCCCA CCATAAAAAC ATTTTTGAGG AGATATGTCC
Betvu;CycA2(2)  ctttcaacta tctgtaccca ccataaaaac attttgagg agatatgttc

               301                                                 350
Betvu;CycA2     ATGCAGCTCA AGCTACT... GAGGATAGTT TGGTTGACTT AGAGTTCTTG
Betvu;CycA2(1)  ATGCAGCTCA AGCTACTTAC GAGGATAGTT TGGTTGACTT AGAGTTCTTG
Betvu;CycA2(2)  atgcagctca agctacttac gaggatggaa aggccgactt agagttcttg

               351                 380
Betvu;CycA2     GCCAAGTACC TAGTA..... ..........
Betvu;CycA2(1)  GCCAAGTATT TAGTGGAACT C.........
Betvu;CycA2(2)  gcaaagtact tcatggagct cggatcctgc
```

CHAPTER 3 GENOMIC LIBRARY CONSTRUCTION AND SCREENING

The isolation of the two partial cyclin sequences described in the previous chapter came as a continuation of the work carried out in our laboratory on the isolation of sugar beet cell cycle related genes (M. Kirby, 1996). The strategy was to use PCR methods and gene specific primers to amplify sequences homologous to known cell cycle related genes. A RACE PCR approach was adopted in order to obtain the full-length cDNAs. However, the extensions towards the 5′ and 3′ ends of the genes gave ambiguous results and the complete cDNA sequences were not obtained. The construction and screening of a sugar beet genomic library was initiated as an alternative approach to obtaining the genes of *Bvcrk1*, *Bvcrk2* and the sugar beet cyclin-Betvu;*CycA2*. In particular, this strategy would permit the isolation of the promoter regions of the genes in question.

3.1 Materials

3.1.1 Plant material

Diploid (2n = 18) sugar beet plants, cultivar Roberta (KWS).

3.1.2 Cloning vectors and bacterial strains

pGEM-T (Promega); Lambda FIX II replacement vector (Stratagene); *E.coli* XL1 Blue, *E.coli* XL1 Blue MRA, *E.coli* XL1 Blue MRA (P2) and VCS 257 strains from Stratagene were maintained and cultivated according to the general procedures described by Sambrook *et al.*, 1989 and following the recommendations of the suppliers.

3.1.3 Kits used for the construction and screening of the genomic library

Lambda FIX II / Xho I partially filled-in vector arm kit (Stratagene); Wizard miniprep and midiprep plasmid purification system (Promega); ChromaSpin Columns: TE1000 (Clontech); NucTrap Probe Purification Columns (Stratagene); Hybond N hybridisation membrane (Amersham); [α^{32}-P]- dCTP Redivue (Amersham); Hyperfilm-MP (Amersham); Glogos Photoluminescent stickers (Stratagene); Klenow Fill-in Kit (Stratagene); Gigapack Gold II packaging extracts (Stratagene); Prime-It Rmt random priming Kit (Stratagene); Geneclean *II*

(Bio 101 Inc.); QIAGEN tip 100 (Qiagen); Wizard Lambda preps DNA purification system (Promega)

3.1.4 Restriction and modifying enzymes, media and solutions

The enzymes used were purchased from Promega, UK unless stated otherwise in the text.

The composition of the media and solutions is given in the Appendix.

3.2 Genomic library construction

The sugar beet genomic library was made with the Lambda FIX II replacement vector (Stratagene). The vector was chosen mainly for two reasons: 1) The Lambda FIX II system takes advantage of *spi* (sensitive to P_2 inhibition) selection which provides a very low background of non-recombinant phages; 2) The arrangement of the Lambda FIX II polylinker permits the isolation of the insert with flanking T3 and T7 promoters as an intact cassette by restriction with the rare cutting enzyme Not I. T3 and T7 promoters flanking the insertion sites can be used to synthesise end-specific RNA probes, and can also be used as primers for PCR.

A summarised procedure for the steps involved in the cloning of gDNA into Lambda FIX II / Xho I partially filled-in vector arms is given below:

1) Isolation of high molecular weight DNA

2) Partial restriction of the DNA with Sau 3A I enzyme in order to obtain the maximum amount of fragments in the size range of 9 to 23 kb.

3) Partial fill-in of the first two nucleotides in the Sau 3A I GATC 5' overhang. This prevents the self-ligation of the fragments and also makes the ends compatible with the vector arm ends.

4) Ligation of the gDNA fragments to Lambda FIX II vector arms.

5) Packaging of the recombinant λ DNA into phage particles and infection of an appropriate host strain, i.e. *E.coli* -XL1 Blue MRA (P2) which will allow the growth of recombinant phages only.

3.2.1 Plant maintenance and collection of samples for DNA isolation

Diploid (2n = 18) sugar beet plants, cultivar Roberta were used for the DNA isolation.

Plants were grown from seedlings in pots in greenhouse conditions. Leaves were collected from 2 to 4 week old plants, which were kept in the dark for 2 days prior to the collection of leaves. The cut leaves were washed with cold sterile water and then left at 4°C for two hours. The midribs and the leaf stems were removed and the leaf material was frozen in liquid nitrogen, and kept at -70°C as 50 g samples.

3.2.2 Plant DNA isolation

High molecular weight DNA was isolated by the method described by Chung *et al.* with minor alterations:

1. 30 g frozen leaf tissue was ground into a fine powder in liquid nitrogen using a mortar and pestle.

2. The powder was transferred to 2L beaker containing 500 ml extraction buffer (50 mM Tris-HCl, pH 8, 100mM NaCl, 50 mM EDTA) and 1% 2-mercaptoethanol and 2% SDS (final volume).

3. The slurry was incubated at 45°C for 45 minutes and then left at room temperature for 15 minutes.

4. An equal volume of phenol was added to the slurry and mixed with gentle rotation by hand.

5. The mixture was poured into 250 ml Beckman centrifuge tubes and spun at 8 000 rpm in a Beckman J14 rotor at 4°C for 10 minutes.

6. The aqueous phase was collected using a wide-tip pipette and DNA was precipitated with 2 volumes of cold 100% ethanol at -20°C for 30 minutes.

7. The precipitated DNA was recovered from the solution with a sterile glass rod. The DNA was washed from the rod into 10 ml 70% ethanol. The tube with the DNA was centrifuged for 1 min at 3000 rpm and the pellet was left to air-dry for 10 min.

8. The pellet was dissolved in 5 ml TE buffer containing 50 $\mu g.\mu l^{-1}$ RNase A and incubated at 37°C for 2 hours.

9. DNA was precipitated by adding 0.1 volume of 3 M sodium acetate, pH 5.2 and 2 volumes of cold ethanol at -20°C for 1 hour.

10. The DNA was pelleted again by centrifugation at 10,000 rpm at 4°C for 10 minutes, washed with 70% ethanol and air-dried for 3 minutes.

11. Finally, the DNA was dissolved in TE buffer, pH 8.0 to a final concentration of 1.2 - 1.5 $\mu g.\mu l^{-1}$ and stored at 4°C until used.

The quality of the DNA was checked by measuring the A_{260} /A_{280} (1.91) and A_{260} /A_{230} (1.78) ratios and by running 0.3% agarose gel of the intact DNA along with lambda DNA (Promega) as a control.

On the 0.3% agarose gels, the sugar beet DNA migrated slower than the lambda control of 49.5 kb (see figure 3.1). The size and purity of the DNA were sufficient to use it as a starting material for the construction of the sugar beet genomic library.

First, aliquots of the DNA were taken for trial partial digests with Sau 3A I and when the optimal concentration of the restriction enzyme was determined, a large scale digest was performed generating the pool of fragments to clone into Lambda FIX II replacement vector (Stratagene).

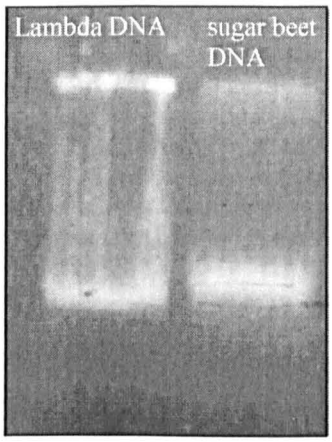


Figure 3.1. 0.3% agarose gel of intact Lambda DNA (lane 1) and sugar beet genomic DNA (lane 2).

3.2.3 Partial digest of genomic DNA

A number of digests were set up in order to determine the optimal amount of Sau 3A I enzyme necessary to produce the maximum number of DNA fragments with sizes from 9 kb to 23 kb.

Sau 3 A I Partial digest of genomic DNA

components	amounts
DNA	20 µg
Sau 3A I buffer, 10X	1x
Sau 3A I enzyme	0.1, 0.05, 0.025, 0.01, 0.008, 0.004 0.0034, and 0.002 U per µg of DNA
d H_2O	up to 25 µl

The digests were incubated at 37°C for 30 minutes. The enzyme was inactivated by adding 0.2 M EDTA, pH 8.0 to a final concentration of 5mM and the digests were loaded on a 0.5% 1x TAE agarose gel along with undigested genomic DNA (0.5 µg) and lambda single cut mixture marker DNA (Sigma). Figure 3.2 shows the ethidium bromide stained gel of a trial partial digest.

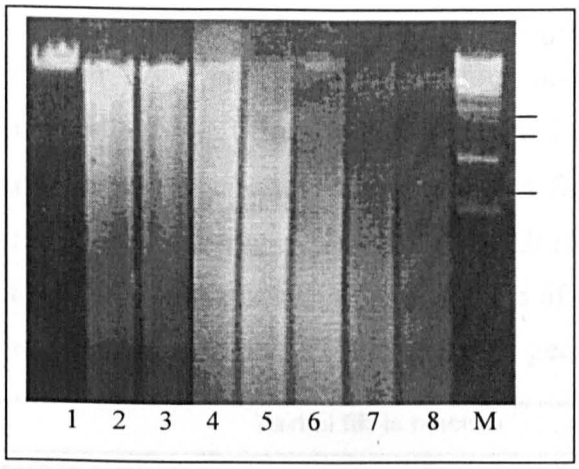


Figure 3.2 Trial partial digest of sugar beet DNA.

Picture of 0.8% 1 x TAE agarose gel. Undigested genomic DNA (lane 1), lanes 2 to 8 Sau 3A I partially digested DNA: 0.002 U (2), 0.0034 U (3), 0.004 U (4), 0.008 U (5), 0.02 U (6), 0.05 U (7), 0.1 U (8) of Sau 3A I enzyme per microgram DNA respectively. M – DNA size markers, size given in kb. From top to bottom the bars represent 48.5, 15.1, 12.2, 10.1 and 8.6 kb.

The optimal concentration of Sau 3A I per μg DNA was determined as the one that produced the maximum ethidium bromide fluorescence from the digested DNA in the range of 9 kb to 23 kb when the gel was visualised under UV light. Using the optimal amount of Sau 3A I enzyme (0.0034 U μg^{-1} DNA), 150 μg of DNA was partially digested but instead of scaling up the reaction, separate digests repeating the trial digest conditions were performed. 10 μl of each digest were loaded onto a 0.3% 1xTAE agarose gel along with the same markers and electrophoresed overnight.

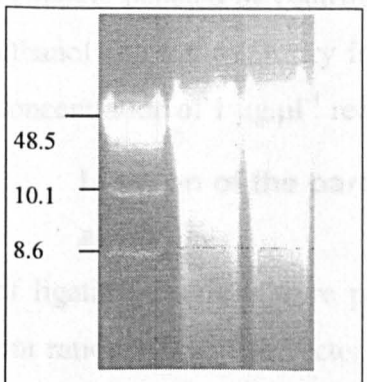


Figure 3.3. Sau 3A I partial digest of sugar beet DNA.
M- marker DNA, size shown in kb; lanes 1 and 2
0.0034U enzyme per μg DNA

The gel was then stained for 15 minutes in 1xTAE buffer containing 5 $\mu g.ml^{-1}$ ethidium bromide, washed briefly in 1xTAE and then visualised under UV light (see figure 3.3). To the remaining 15 μl from each digest 0.2 M EDTA, pH 8.0 was added to final concentration of 5 mM. The DNA was purified through CHROMA SPIN™-TE 1000 columns (Clontech) and taken for the partial fill-in reaction.

3.2.4 Partial fill-in of Sau 3A I digested DNA

The Sau 3A I enzyme cuts the four nucleotide site GATC, leaving 5'-CTAG overhangs incompatible with the CT overhangs of the Lambda FIX II Xho I vector arms. In order to achieve the ligation of the Sau 3AI digested DNA fragments to the vector arms, a partial fill-in reaction was carried out using a Klenow Fill-in Kit (Stratagene) following the suppliers instructions. The fill-in reaction also prevented the ligation of chimeric genomic fragments to the vector arms and ensured that the resulting clones from the genomic library did not contain artefacts.

Partial fill-in reaction	
reaction components	**final quantities in the reaction**
DNA	50 µg
fill-in buffer,10x	30 µl
dATP (10 mM)	5 µl
dGTP (10 mM)	5 µl
Klenow DNA polymerase (5 U μl^{-1})	3 µl
d H_2O	to final volume 300 µl.

The reaction was incubated at 20°C for 15 minutes. It was stopped with addition of 0.2 M EDTA, pH 8.0 to final concentration of 5 mM.

The DNA from the fill-in reaction was purified through a Chroma spin™-TE 1000 column (Clontech), precipitated with 0.3 volumes of 3 M sodium acetate, pH 5.2 and 2 volumes of cold 100% ethanol, pelleted by centrifugation at 13 000 rpm at 4°C for 5 minutes, washed twice with 70% ethanol and left to air dry for 2 min. The DNA was redisolved in sterile distilled water to final concentration of 1 $\mu g.\mu l^{-1}$ required for the ligation into the Lambda FIX II vector arms.

3.2.5 Ligation of the partially filled-in DNA fragments to Lambda FIX II vector arms

Sets of ligation reactions were performed with T4 DNA ligase (New England Biolabs) using different ratios of insert to vector DNA. The concentration of vector arms was kept at 1 µg per ligation and the concentration of insert DNA was varied from 0.5 to 1.5 µg per reaction:

Ligation reaction	
reaction components	**final quantities in the reaction**
vector	1 µg
insert	0.5 to 1.5 µg (see table 3.1)
ligase buffer	1 x
T4 DNA ligase	4.8 U
dH_2O	to 5 µl

Two control reactions were included in order to estimate the efficiency of ligation and the background religation of the vector to itself.

	Positive control	Negative control
components		
vector	1 μg	1 μg
pMF (12 kb) test insert	3 μg	not included
ligation buffer	1 x	1 x
T4 DNA ligase	3 U	3U
d H_2O	to 5 μl	to 5 μl

The ligation reactions were carried out at 4°C for 16 hours, the ligase was then heat inactivated at 75°C for 15 minutes. The ligations were taken for the next step in the cloning protocol - packaging of the recombinant lambda DNA into viable phage particles.

3.2.6 *In vitro* packaging of the recombinant lambda DNA and titration of the genomic library

The ligated insert/lambda vector arms were packaged into lambda particles using Gigapack Gold II packaging extracts (Stratagene) following the recommendations of the supplier.

Different amounts from the ligation reactions were packaged and the packaging efficiency per μg packaged DNA was calculated using the equation:

$[N \times d \times Vtotal] \div [Ptotal \times Pplated]$, where:

N is the number of plaques counted; d is the dilution factor; $Vtotal$ is the total packaging volume; $Ptotal$ is the total μg of packaged DNA; and $Pplated$ is the number of μl plated.

Controls were included as follows: to test the efficiency of the packaging reaction wild type lambda (λ c*I* 857 Sam) provided with the kit was packaged and titered in *E.coli* VCS 257 host strain. Positive and negative control ligations were also packaged (table 3.1).

The packaged lambda was used to infect the appropriate host bacterial strains and titers of the libraries were determined (table 3.1) by plating 10^{-2} and 10^{-3} dilutions of the original libraries (samples **A**, **B**, **D**, **E**, **G**, and the controls).

The bacterial strains used were grown in LB media supplemented with 10 mM $MgSO_4$ and 0.2% maltose, to OD_{600} = 0.8 and collected by centrifugation at 2000 rpm for 10 minutes. The bacteria were resuspended in 10 mM $MgSO_4$ to OD_{600} of 0.5. An aliquot of 200 μl was infected with 10 μl of 10^{-2} and 10^{-3} dilutions of the original packaging reactions (500 μl). The infected bacteria were

mixed with 3 ml LB top agar (50 °C) and poured onto 90 mm NZY plates. The number of plaques per plate was counted after 12 to 14 hours of incubation at 37°C.

The titer of the library was calculated as the number of plaques counted x dilution factor x total packaging volume / number of µl plated.

Table 3.1. A, B, C, D, E and G - different packaging reactions (total volume 500 µl). Libraries D and G were used for amplification and screening respectively.

sample	A	B	C	D	E	G	positive control	negative control	λDNA control
λ DNA arms, µg	1	1	1	1	1	1	1	1	N /A
insert DNA, µg	1	0.5	2	1.5	0.75	0.75	0.3	-	N /A
volume of ligation packaged,µl	1 0.4 µg	1.5 0.45µg	-	1.5 0.75µg	2.5 0.88µg	4 1.4µg	2 0.52 µg	2.5 0.5 µg	1 0.2µg
efficiency μg^{-1}	5.1x 10^5	1.3x 10^6	-	7.5x 10^5	1.2x 10^6	1.28x 10^6	1.25x 10^6	-	1.2x 10^8
titer	2.53x 10^5	6.1x 10^5	-	6.08x 10^5	9.8x 10^5	1.67x 10^6	5.88x 10^5	-	2.4x 10^8

3.2.7 Estimation of the average size of the cloned genomic inserts and the quality of the genomic libraries

Six random plaques from each library were taken for amplification. Lambda DNA was isolated from each of the amplified plaques and digested with Not I, Sal I, Hind III restriction enzymes. Restrictions were in a volume of 10 to 20 µl as follows:

Random digests of Lambda genomic clones	
reaction components	
DNA	0.2 to 0.8 µg
enzyme buffer 10 x	1 x
restriction enzyme (Not I, Sal I or Hind III)	2.5 U
d H_2O	up to final volume of 10 - 20 µl

The digests were left at 37°C for 4 hours, mixed with loading dye (5x) and loaded on 0.6% 1x TAE agarose gel along with DNA size markers. The gel was run at 25 V cm^{-1} overnight then the DNA was visualised with UV light and the approximate sizes of the inserts determined by comparison with the DNA size markers migration in the gel.

According to the sizes of the DNA bands determined from the agarose gel, the average size of the inserts was between 14 kb-16 kb. The smallest insert appeared to be 10.5 kb and the largest one was slightly above 21 kb.

Given the size of the sugar beet haploid genome - 750 Mbp (Arumuganathan and Earl, 1991), the number of plaques needed to clone a unique insert (15 kb) with a probability P= 0.98 would be 1.96×10^5.

The equation used for the calculation was: $N = 1 - \ln(1-P) \div \ln(1-f)$,

where N is the number of plaques needed; P is the probability (0.98); f represents the average size of the cloned fragments (15 kb) as a portion of the length of the haploid genome (750 Mbp for sugar beet).

The titer of sample **A** (2.5×10^5) is close to this figure, these of samples **B** and **D** are about three times higher, sample **E** has about 4.5 times more pfu, and the titer of sample **G** is almost 10 times the minimum pfu required for the genomic library to be representative (table 3.1).

The titering of the recombinant phages was done using *E.coli* XL1 Blue MRA (P2) host cells for inoculation which allowed only recombinant phages to grow. The control for religation of lambda arms to the stuffer fragment (negative control in table 3.1) did not grow in this bacterial strain indicating that the titers of the libraries did not have any background due to the presence of non-recombinant phages.

3.2.8 Amplification of the genomic library

One of the libraries, sample **E**, was amplified as follows: ten 25.5 µl aliquots (a total of 4.5×10^5 pfu) of the packaged library were used to infect 600 µl of *E. coli* XL1 - Blue MRA (P2) cells of $OD_{600} = 0.5$ in 10 mM $MgSO_4$. Infected cells were mixed with 6 ml of LB top agar, poured onto 137-mm NZY plates and left to grow for 10 - 12 hours at 37°C.

Next, 8 ml of SM buffer was added to each plate and the plates were left overnight at 4°C with occasional gentle shaking.

The SM buffer with the diffused phage particles was collected as thoroughly as possible and another 2 ml of fresh SM buffer were used to wash the surface of the plates. The SM buffers were pooled together and mixed with chloroform (final concentration 5%). After 15 minutes incubation at room temperature, the cell debris was pelleted by centrifugation at 500 x g for 10 minutes. The supernatant was transferred to a sterile 50 ml tube and chloroform was added to

0.3% concentration to prevent the growth of bacteria. The amplified library was stored at 4°C in the dark.

The titer of the amplified library was determined using *E. coli* XL1 - Blue MRA (P2) host cells. Plaques were counted after 10 hours of incubation and the titer of the library calculated using the equation :

$$Titer = \frac{Np \times d \times Vol}{P}$$

Titer is measured in plaque forming units per ml; *Np* is number of plaques counted; *Vol* is the total volume of the library (500µl); *P* is µl of pfu plated.

3.2.9 Lambda phage amplification

Lambda phages from the library were amplified by either liquid culture or plate lysate methods. The resulting lysates were stored at 4°C under 0.3% chloroform until used for DNA isolation.

The preparation of **liquid lysates** followed the protocol of Sambrook *et al.* (Molecular cloning Vol2, 1989) for infection at high multiplicity.

1. 10 ml of LB cultures supplemented with 0.2% maltose and 10mM $MgSO_4$ were inoculated with a single colony of the host bacterial strain, *E.coli* XL1 MRA and grown overnight.

2. 100 µl of the overnight bacterial culture was transferred to 50 ml of fresh LB containing 0.2% maltose plus 10mM $MgSO_4$ and grown until it reached OD_{600} = 0.5 (3 to 5 hours).

3. Each 50 ml culture was inoculated with 10^6 to 10^{10} pfu of lambda phage particles and grown at 37°C until lysis occurs, normally about 4 to 6 hours.

4. To stop the growth of the remaining bacteria, 1 ml chloroform was added to the 50 ml cultures and left for 15 minutes at 37°C.

5. The lysates were then centrifuged at 8 000 rpm for 10 minutes. The clear lysates were transferred to new 50 ml tubes and stored at 4°C until used for DNA isolation or as inoculum for further amplification.

The **plate lysates** were prepared by a protocol similar to the one suggested by Stratagene for Lambda FIX library amplification, with alteration to the volumes used.

1. 10 ml LB culture was inoculated with a single colony from the host bacterial strain *E. coli* XL 1 MRA or *E.coli* XL1 MRA (P2) and left in a 37°C shaker incubator to grow overnight.

2. A fresh LB supplemented with 0.2% maltose and 10mM $MgSO_4$ was inoculated with 50 µl of the overnight bacterial culture and left to grow until it reached OD_{600} = 0.8.

3. The bacteria were collected by centrifugation for 10 minutes at 4°C at 2 000 rpm and the LB media was removed as thoroughly as possible.

4. The bacteria were resuspended in ice cold 10mM $MgSO_4$ to half of the original volume of the culture and kept on ice or stored at 4°C overnight.

5. Immediately before the inoculation with phage particles, the bacteria were diluted with 10mM. $MgSO_4$ to $OD_{600} = 500$.

6. 200µl or 600µl of bacterial cells of $OD_{600} = 500$ were inoculated with 10^3 or $5x10^3$ respectively and left for 20 minutes at 37°C.

7. 0.5% melted agarose (48 - 50°C) was mixed with the bacteria and the agarose was poured over LB-agarose petri dishes preheated to 37°C. The plates were left overnight at 37°C.

8. 3 ml (90 mm petri dish) or 8 ml (137 mm petri dish) of SM buffer was added to the plate lysates. Plates were left overnight in a cold room on a slow moving shaker.

9. The SM buffer containing the phages was collected from the petri dishes as thoroughly as possible and chloroform was added to final volume 5%. The tubes containing the lysates were vortexed briefly and then centrifuged at 3 000 rpm for 15 minutes.

10. The clear lysates were transferred to new tubes and were kept with 0.3% chloroform until used for DNA isolation or as an inoculate for infecting bacteria.

3.2.10 Lambda DNA isolation

Lambda DNA was isolated from bacterial lysates using several methods depending on the amount of DNA needed.

Lambda DNA taken to estimate the average size of the inserts in the library (3.2.7) was isolated using a mini-prep method (Promega). In the rest of the applications the DNA was isolated using a QIAGEN-tip 100 kit (Qiagen) or Wizard Lambda DNA purification system (Promega) and following the instructions from the suppliers.

- Mini-prep isolation of lambda DNA

1. 100µg/ml DNase and RNase were added to each lambda lysate. The lysates were incubated for 30 minutes at 37°C.

2. An equal volume of phage precipitation solution (20% PEG 8000, 2M NaCl) was added to the lysates and they were left for 1 hour on ice.

3. The precipitated phage particles were recovered by centrifugation at 10 000g for 20 minutes at 4°C.

4. The supernatant was removed with a pipette and the precipitate was drained.

5. 1 ml of phage buffer (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10 mM $MgSO_4$) was added to the pellet per 10 ml of the starting lysate volume. The phage particles were resuspended by gentle vortexing.

6. The debris was removed by centrifugation for 2 minutes at 8 000g at 4°C.

7. The supernatant was extracted with 1 volume TE saturated phenol/chloroform (1:1) by vortexing for 1 minute and then centrifugation for 5 minutes at 13 000 rpm.

8. The aqueous phase was extracted with 1 volume TEN saturated phenol/chloroform, again vortexed for 1 minute and centrifuged for 5 minutes at 13 000 rpm in a bench top centrifuge.

9. The aqueous phase was extracted once with 1 volume chloroform:isoamylalcohol (24:1), vortexed again for 1 minutes and centrifuged for 5 minutes at 13 000 rpm.

10. The aqueous phase was precipitated with 1 volume of iso-propanol, left at -70°C for 20 minutes and then centrifuged at 13 000 rpm for 10 minutes at 4°C.

11. The pellet was washed with 1 ml 70% ethanol, dried under vacuum and resuspended in 50 to 100 µl of TE buffer. The lambda DNA was stored at 4°C.

3.3 Sugar beet genomic library screening

The screening of the library generally involved the taking of plaque lifts from high density plated phages and hybridisation of the membranes to the appropriate labelled probes. Once the positive signals had been detected, a second round of plaque lifts and hybridisations was performed to isolate single positive plaques.

3.3.1 Probes used for the screening of the library

1. *Bvcrk1* - a PCR amplified 604 bp Nco I / Sal I fragment (EMBL accession number Z71701) in pGEM-T (Promega) isolated from sugar beet genomic DNA from cell suspension cultures , cultivar Regina (3n) by M. Kirby (PhD Thesis, 1996) using primers with homology to the *cdc2* gene subdomain I - GEGTYG and subdomain VII - WYRAPE (Martinez *et al.*, 1992). The sequence of the fragment is given in figure 4.2, chapter 4.

2. Three cyclin-like sequences from sugar beet were taken as probes for screening. They were nearly identical and were used together as a mix in the hybridisation solution. The sequence of the fragments and the details of how they were isolated is given in chapter 3 of this thesis.

3. *Bvcrk2* - a 380 bp fragment (EMBL accession number Z71704) cloned into pGEM-T plasmid vector (Promega). The sequence was isolated by M. Fowler from root tip total RNA using primers homologous to the GEGTYG (Martinez *et al.*,1992) and HRDLKPQ (Hanks *et al.*,1988) regions in cdc2 gene. The fragment shares homology with the *Arabidopsis* MHK gene and belongs to the MAK gene subgroup (Matsushime *et al.*, 1990).

3.3.2 Maintenance of the clones and probes labelling

Glycerol stocks of the plasmids were kept at -20°C. Plasmids were isolated from 10 ml overnight cultures of the bacteria using the Promega Wizard miniprep kit and following the instructions supplied with it. The *crk1*, *crk2* and cyclin inserts were cut out from the MCS of the plasmids with appropriate restriction enzymes and the resulting bands purified from the agarose gel by Geneclean II kit according to the instructions of the manufacturer (Bio 101, Inc.). The quality and quantity of the purified fragments was checked on 1% agarose gels. The inserts were stored frozen at -20°C until needed for labelling.

The *crk1*, cyclins and *crk2* probes were labelled with [α^{32}P]- dCTP using Prime-It®Rmt random priming kit (Stratagene).

50 ng of "genecleaned" probe was used per labelling reaction. The labelled probe was purified through NucTrap®Probe Purification Columns (Stratagene). In both cases the manufacturer's instructions were followed. The probes were used immediately for hybridisation and the time of the labelling was calculated so that the end of the labelling reaction coincided with the end of the pre-hybridisation step.

3.3.3 Plaque lifts

3 x 10^5 plaques from sugar beet genomic library **G** were screened with each probe. This number of plaque forming units covered one genome size of sugar beet of 2.99 x 10^5. An average of 3.75 x 10^4 pfu (for *crk1* and *crk2*) or 5.0 x 10^4 (*Bvcyc*) per 600 µl *E. coli* XL1 -Blue MRA (P2) of OD_{600} = 0.5 in 10 mM $MgSO_4$ were used per plate. The plaques were left to grow for 12 hours at 37°C, then left for 3 hours at 4°C in order to prevent the top agarose from sticking to the membranes. Plaque lifts were taken using 134mm circular Hybond™-N membranes (Amersham International plc, UK) following the recommendations of the supplier. The membranes were left to dry and then baked for 2 hours at 80°C to fix the DNA.

3.3.4 Hybridisation

Hybridisations were carried out at 42°C for 16 hours in 15 ml hybridisation solution with 50 ng per membrane. Each hybridisation bottle contained three membranes separated by hybridisation mesh (Hybaid) to reduce the background signal. The prehybridisation step was at 42°C for 4 hours. The prehybridisation solution (2 x PIPES, 50% deionised formamide, 5 x Denhardt's, denatured sonicated salmon sperm 100 $\mu g\ \mu l^{-1}$) was exchanged with fresh hybridisation solution (same as prehybridisation solution) prewarmed to 42°C. The radioactive probe was denatured for 5 minutes at 95°C and added to the hybridisation solution..

The membranes were washed in plastic boxes with a prewarmed solution of 0.1x SSC and 0.1% SDS (55°C) until the radiation on the blank filter included along with the other filters in the hybridisation bottle dropped close to the background level of radioactivity of 2 Bq cm^{-2}. The membranes were wrapped in Cling Film and exposed to Hyperfilm™-MP (Amersham) X-ray film for 24 hours to 7 days at room temperature. The filters were oriented using photo-luminescent stickers Glogos™ (Stratagene). The films were aligned with the plaques, the positive signal areas were excised with a scalpel blade and placed in 1 ml SM buffer. Plaques were left to diffuse overnight at 4°C and their titres were checked by infecting the host bacterial strain as described earlier in this chapter.

The positive plaques from the first screen for with *crk1* were rescreened at a plaque density of 50-60 plaques per 90 mm plate which allowed for the isolation of single positives. Fifty ml bacterial cultures of OD_{600}=0.5 were infected with the single positives at high multiplicity (Sambrook *et al.*, 1989) for lambda DNA isolation.

3.3.5 Screening results

According to the intensity of the signal detected the plaques were divided into four categories: very strong signal (+++), strong signal (++), weak signal (+), and very weak signal (+/-). Table 3.2 gives a summary of the results.

Twenty nine positive signals were detected with crk1 probe. The plaques with very weak signal (+/-) were excluded from the second screening, leaving 16 plaques to test. After the rescreening the number of positives was reduced to 6, of which three had strong to very strong signal the rest gave weak signals.

In the following screenings with crk2 and cyclin probes only the first three categories of signals (+++; ++; +) were recorded (see table 3.2) and only the plaques with very strong and strong signal were selected for rescreening. The plaques isolated after the first screening for cyclin and

crk2 genes were stored as phage plugs in SM buffer and were available for further characterisation in the laboratory.

Table 3.2 **Summary of the screening results.** An initial screening was carried out for three different genes. Positives were selected for second screening. The resulting single plaques were taken for further characterisation in the case of crk1.

probe used for screening	total number of plaques screened	plaques plated per plate	number of positives after the first screening	positives taken for the second screening	number of positives after the second screening
crk1	3×10^5	3.75×10^4	29	16 (+++, ++, +)	6
cyclin	3×10^5	5×10^4	12	5 (+++, ++)	N/A
crk2	3×10^5	3.75×10^4	18	9 (+++, ++)	N/A

CHAPTER 4 THE *Bvcrk1* GENE: SEQUENCE, INTRON-EXON ORGANISATION AND THE PREDICTED CRK1 PROTEIN

This chapter describes the characterisation and sequencing of the *crk1* Lambda Fix II clones obtained after the screening of the sugar beet genomic library (chapter 3).

The presence of *crk1* homologous sequences in the lambda clones was confirmed by Southern blot analysis. As a result, two *crk1* clones were defined. About 10 kb of each of the two genomic clones was sequenced. The information from the genomic sequence of the *crk1* gene allowed for the rapid isolation of its the full-length cDNA sequence.

4.1 Materials

4.1.1 Cloning vectors and bacterial strains

The plasmids pGEM-T (Promega) and pZErO-2 (Invitrogen) were used for cloning of PCR and restriction digest products.

E.coli XL1 Blue and *E.coli* XL1 Blue MRA strains from Stratagene and *E.coli* TOP10F' from Invitrogen were maintained and cultivated according to the general procedures described by Sambrook *et al.*, 1989 and following the recommendations of the suppliers.

4.1.2 Primers

Primers were either synthesised in the laboratory on a PCR MATE DNA Synthesiser or ordered from a commercial supplier (Cruachem or Gibco). For the sequence of the oligonucleotides refer to the primer tables in the Appendix.

4.1.3 Kits and special reagents

ChromaSpin Columns: TE1000, DEPC-200, H2O-200 (Clontech); Hybond N hybridisation membrane (Amersham); [α^{35} -S]- dATP, Redivue (Amersham); [α^{33}-P]- ddNTPs, Redivue labelling mix (Amersham); Hyprefilm-MP and Hyperfilm-βMax (Amersham); Kodak BioMax X-ray film; Sequenase Version 2.0 and ThermoSequenase radiolabeled terminator cycle sequencing kit (Stratagene); GeneAmp XL PCR Kit (Perkin Elmer); NEBlot Phototope Kit (New England Biolabs); Geneclean *II* (Bio 101 Inc.); QIAGEN tip 100 (Qiagen); Wizard Lambda preps DNA purification system (Promega)

4.1.4 Restriction and modifying enzymes, media and solutions

The enzymes used were purchased from Promega, UK unless stated otherwise in the text.

The composition of the media and solutions is given in the Appendix.

4.2 Methods

4.2.1 Characterisation of the *crk1* positive plaques

Six *crk1* positive clones were isolated from the genomic library (described in chapter 3). They had to be characterised further by PCR and Southern blot analysis in order to narrow down the number of clones for subcloning and sequencing.

Figure 4.1 gives a summary of the screening and characterisation process.

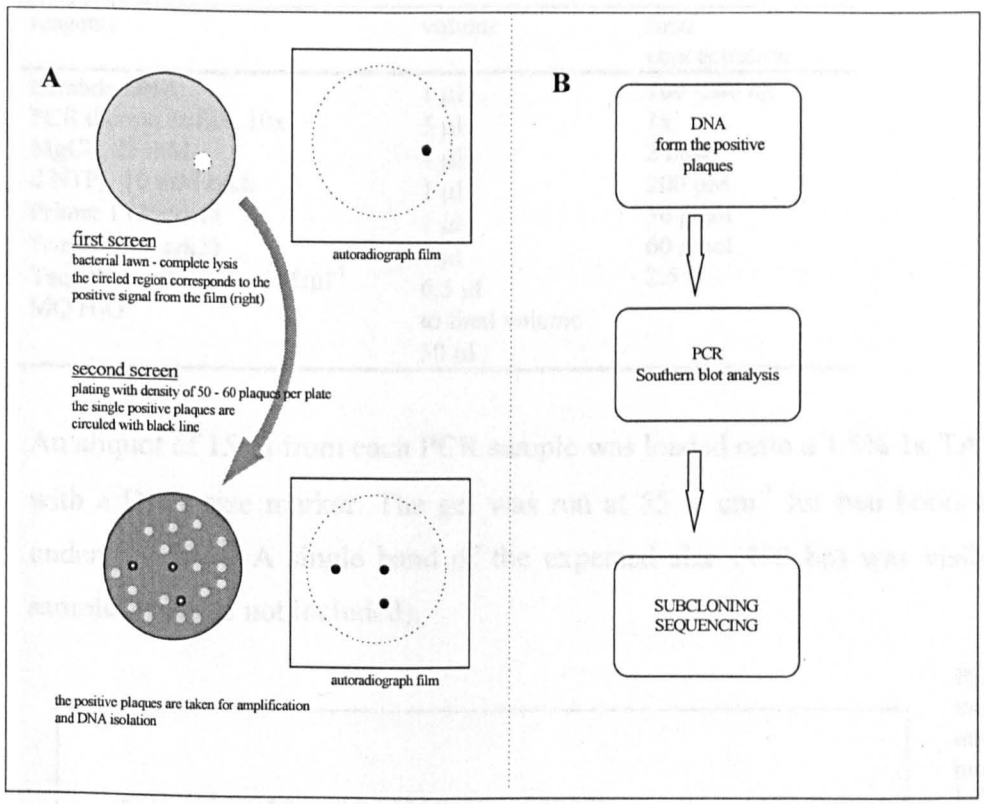


Figure 4.1. Genomic library screening (A); plaque characterisation (B)

In part A the screening of the library is shown. The grey circles represent plaque plates. The part of the bacterial lawn, which aligned with the autoradiograph film signal, was taken to be propagated and rescreened. The plaques isolated after the second screening as positive were taken for further work: (B) In order to narrow down the number of true positives and to select for the homologue of the *crk1* probe, the DNA from the positive plaques was used in PCR and Southern blot analysis. The resulting positive clones were subcloned and sequenced.

5.2.1.2 PCR

Initially, the presence of *crk1* homologous sequence in the isolated plaques was confirmed by PCR. Two internal *crk1* specific primers: 3′-crk1 and 5′-crk2 (the two primers are represented in blue colour on figure 5.2) were used to amplify a region of about 400 bp which is present on the known part of the *crk1* gene.

PCR reaction conditions: an initial denaturation step of 4 minutes at 94°C was followed by 30 cycles of 1 minute at 94°C, 2 minutes at 55°C, 1 minute at 72°C; final extension step (72°C) of 10 minutes.

PCR contents		
reagents	volume	final concentration
Lambda DNA	1 µl	100 -200 ng
PCR thermo buffer, 10x	5 µl	1x
$MgCl_2$, 25 mM	4 µl	2 mM
d NTP_S, 10 mM each	1 µl	200 µM
Primer 1 (3′ crk1)	1 µl	50 pmol
Primer 2 (5′ crk2)	1 µl	60 pmol
Taq DNA polymerase, $5U\mu l^{-1}$	0.5 µl	2.5 U
MQ H_2O	to final volume 50 µl	

An aliquot of 15 µl from each PCR sample was loaded onto a 1.5% 1x TAE agarose gel along with a DNA size marker. The gel was run at 55 V cm^{-1} for two hours and then visualised under UV light. A single band of the expected size (400 bp) was visible for some of the samples (picture not included).

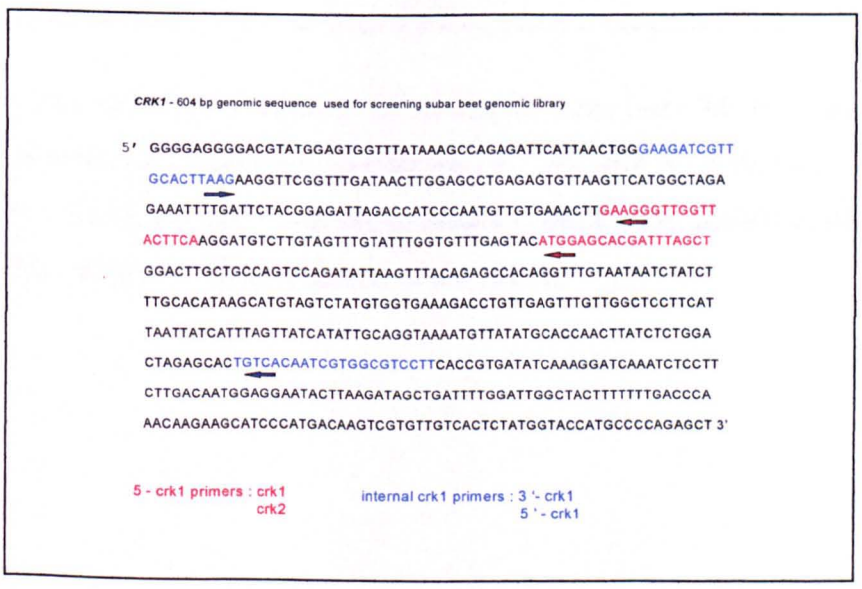


Figure 5.2. Crk1 probe used for screening and orientation of the primers used in the PCR.

The *crk1* PCR probe used for the screening of the sugar beet library. The 5′ specific primers are shown in red and the crk1 internal primer set are in blue colour. The arrows under the sequence show the orientation of the primers.

A second PCR was carried out on the lambda positive clones using the T3/T7 promoter primers from the MCS of the lambda vector in combination with one of the 5′-gene specific primers (primers in red colour in figure 4.2). The gene specific primers were positioned at the 5′end of the known cDNA region of the *crk1* gene. It was expected that when the gene specific primers were used together with a primer from the MCS of the vector, the orientation of the inserts and the approximate distance between the cloning site and the known part of the 5′ end of the gene could be determined.

The GeneAmp® XL PCR Kit (Perkin-Elmer, Roche Molecular Systems, Inc., NJ, USA) was used because it takes advantage of r*Tth* DNA polymerase, XL which ensures correct amplification of large templates (5 kb - 35 kb).

XL PCR reaction conditions : An initial denaturation step at 94°C for 2 minutes was followed by 16 cycles of denaturation at 94°C for 40 seconds, annealing at 50°C for 1 minute, and extension at 70°C for 5 minutes; then 12 more cycles followed but with an elongation step 15 seconds longer for each following cycle; the final extension time at 72°C was 10 minutes.

XL PCR		
reagents	volume	final concentration
DNA template		100 - 200 ng
$Mg(OAc)_2$, 25 mM		1.1 mM
dNTP, 10 mM each	1 µl	200 µM
XL buffer, 3.3 x		1 x
Promoter primer (T3 or T7)		80 pmol
Gene specific primer (sq5′crk1 or sq5′crk2)	1 µl	100 pmol
r*Tth* DNA Polymerase, 4 U μl^{-1}	0.5 µl	2 U
dH_2O	to final volume 50µl	

After the amplification, a 35 µl aliquot from each XL PCR was loaded on a 1% 1 x TAE agarose gel (Boeringer Mannheim UK Ltd.) run for 4 hours at 45 V cm^{-1} (figure 4.3). Table 4.1 shows the combinations of primers used on the lambda positives, the approximate size of the bands and the orientation of the inserts.

sample	Primer sets	Approximate PCR product size, Kb
1.2	T3 / 5′crk 1 T7 / 5′ crk1	
2.2	T3 / 5′crk 1	-
4.6	T3 / 5′crk 1 T7 / 5′ crk1	
6.2	**T3 / 5′crk 1** T7 / 5′crk 1	5 kb -
7.1	T3/ 5′crk 1 **T7/ 5′crk 1**	- band 1 - 4.3 kb **band 2** - 3.3 kb band 3 - < 1.5 kb
8.4	**T3 / 5′crk 1or 5′ crk 2** T7 / 5′crk 1or 5′ crk 2	band 1 - 4.3 kb **band 2** - 3.3 kb band 3 - < 1.5 kb

Table 4.1. Combination of primers and approximate sizes of the products of the XL PCR.

The two gene specific primers 5′ crk1 and 5′ crk2 (red colour on figure 4.1) were used in combination with the promoter primers T3 and T7 from the Lambda FIX II vector; the primer combinations which resulted in PCR amplification products as determined by agarose gel electrophoresis are given in bold. In the case of lambda clones 7.1 and 8.4 several bands were observed and the strongest of the bands is shown here in bold.

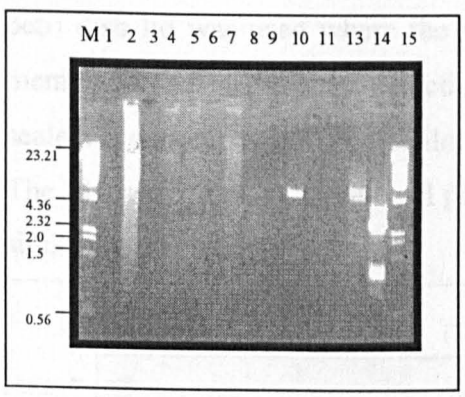


Figure 4.3. XL PCR of the crk1 positive genomic clones in Lambda FIX II: A 1% agarose gel of XL PCR. Lane M and 15 - marker DNA; the size (kb) of the bands is indicated on the left. Lanes 2 to 14 - XL PCR of lambda clones: 1.2 (lane 2), 6.2 (lane 10), 7.1 (lane 13 and 14). The PCR products from 8.4 are not shown on this gel but they had the same banding pattern as 7.1.

4.2.1.2 Southern blot analysis

The six lambda positive clones for *crk1* (1.2, 2.2, 4.6, 6.2, 7.1, 8.4) were digested with Xba I and with Xba I + Bam H I. The digests were left overnight at 37°C and the enzymes were inactivated by incubating at 70°C for 10 minutes.

Digest reaction components	
λ DNA	100 - 500 ng
enzyme(s), 10 U	1 U
restriction buffer, 10x	1 x
MQ H_2O	to final volume 25 µl

The reactions were mixed with loading dye (5x) and loaded on a 0.6% 1 x TAE agarose gel run at 45 V cm^{-1} for five hours.

The gel was visualised under UV light and a photograph of the gel was taken for reference.

The gel was blotted onto Hybond™ - N membrane (Amersham International, UK) following the instructions of the supplier for an overnight capillary blot transfer in 20xSSC, then it was washed briefly in 1 x SSC and air dried for 10 minutes. The DNA was fixed to the membrane by oven baking at 80°C for 2 hours. The hybridisation was carried out using biotin labelled (NetBlot Phototop Kit, New England Biolabs) *crk1* probe(604 bp Nco I /Sal I fragment), following these steps: 1) The membrane was floated briefly in 6 x SSC; 2) Prehybridization at 65°C in 10 ml prehybridization buffer (6 x SSC, 5 x Denhardt's, 0.5% SDS, 100μg ml^{-1} denatured salmon sperm DNA) for 2 hours; 3) The prehybridization solution was changed with 7 ml hybridization solution (same as above) containing the denatured *crk1* probe at a concentration of 20 ng/cm^2 membrane and hybridised overnight at 65°C in a roller bottle (Hybaid); 4) Washings were carried out in 2 x SSC, 0.1% SDS in the hybridisation bottle (Hybaid) twice at room temperature for 5 minutes and twice in 0.1% SSC, 0.1% SDS at 68°C for 20 minutes.

The membrane was taken for chemiluminescent detection following the instructions of the NEBlot™Phototope™ Kit used for the probe labelling. Instead of plastic seal bags, a 137-mm petri dish lid was used where the reagents were added in the same volumes directly to the membrane. After the final detection step, the membrane was placed in a plastic bag and sealed. It was exposed to Hyperfilm MP (Amersham) film for up to 1 hour (figure 4.4).

The film was aligned with the gel picture taken before blotting and the bands giving a positive signal were noted.

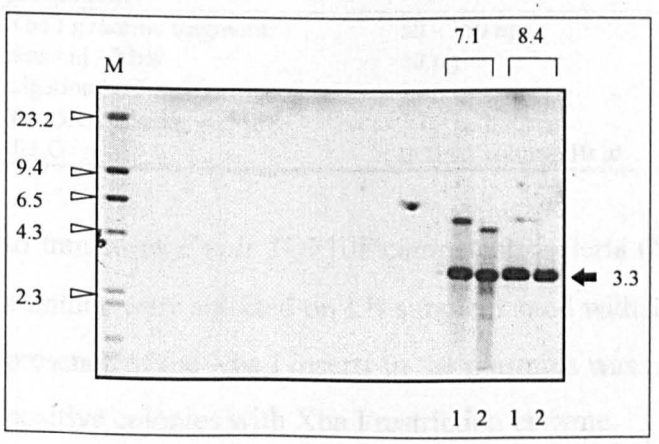


Figure 4.4 **Southern-blot analysis of *crk1* positive lambda clones.**

Xba I (1) and Xba I + Bam H I (2) digested lambda clones 1.2, 2.2, 4.6, 6.2, 7.1 and 8.4. The end 4 lanes are the 7.1 and 8.4 digests respectively. M - biotinylated size marker DNA, size of bands given in kb

4.2.1.3 Cloning of the Xba I bands

A second Xba I digest of the two positive clones 7.1 and 8.4 was done using 1µg DNA, 2.5 U of the enzyme in total volume of 25µl for 2 hours at 37°C. The reaction was heat inactivated at 70°C for 10 minutes and loaded on a 0.8% 1 x TAE agarose gel run at 50 V cm^{-1} until the bands from the digest were separated well enough to be cut out from the gel (figure 4.5).

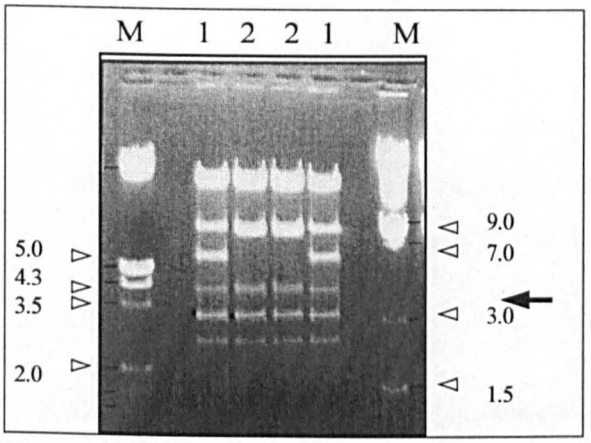


Figure 4.5 Xba I digest of 7.1 and 8.4.

Xba I digest of 7.1 (lanes 1) and 8.4 (lanes 2). M - DNA size markers, the sizes of the bands are shown in kb. The 3.3 kb positive band (figure 4.4) is marked with a dot on the gel picture. The two smallest bands from the digest (1.7 and 0.8 for 7.1 and 0.8 and 0.3 from 8.4) are not visible on this picture.

The bands were cleaned from the agarose with the Geneclean *II* kit (*BIO 101 Inc.*, La Jolla, CA). The purified fragments from the Xba I digest of the Lambda Fix II *crk1* positive clones 7.1 and 8.4 were subcloned into Xba I linearised pZErO-2 plasmid (Invitrogen) in ratios of 1:1 to 1:3 of vector: insert.

Ligation reaction components	
Xba I genomic fragment	50 - 150 ng
plasmid / XbaI	50 ng
Ligation buffer, 10 x	1 x
T 4 DNA Ligase, 6 U μl^{-1}	3 U
dH_2O	to final volume 10 µl

The reactions were incubated at 25°C for two hours and the ligase (New England Biolabs Inc., USA) was heat inactivated at 75°C for 15 minutes.

Half of each ligation (5µl) was used to transform *E.coli* TOP10F'competent bacteria (Nishimura *et al.*) and the positive bacterial colonies were selected on LB supplemented with 2mM IPTG and 25 $\mu g.ml^{-1}$ kanamycin. The presence of the Xba I inserts in the plasmids was checked by digests of the plasmids from the positive colonies with Xba I restriction enzyme.

Table 4.3 gives a summary of the subcloning process with the approximate sizes of the bands and the number of the subclones obtained after transformation together with information about the plasmid subclones which were sequenced.

Table 4.2 Subcloning of Xba I bands from lambda clones 7.1 and 8.4.

The table gives the number and approximate sizes of the Xba I fragments used for cloning in the Xba I site of pZErO-2. GC-genecleaned bands. * For convenience, the bands are arranged in descending order according to their size. ** only some clones were taken for sequencing, and only some of them were completed. ✓ indicates the bands taken for subcloning.

clone	GC bands*	estimated size, kb	number of subclones	names of subclones taken for sequencing **
7.1	1	6	-	-
	2	4.3	-	-
	3	3.3✓	4	4H, 4J
	4	2.7✓	8	7.1-4 (partially sequenced)
	5	1.7✓	0	-
	6	0.8✓	2	7M (1) and 7M (2)
8.4	1	9	-	-
	2	4.4 - 4.5✓	8	-
	3	3.3✓	5	8.4-**1** and 8.4-**2**
	4	2.7✓	24	8.4-4(partially sequenced)
	5	0.8✓	6	25 and 27
	6	0.3✓	1	8.4-**0.3**

4.2.1.4 Sequencing of the *crk1* plasmid subclones and lambda clones

The plasmid subclones of 7.1 and 8.4 were sequenced with Sequenase Version 2.0 kit from Amersham. The labelling followed the standard procedures described by the supplier. The samples were run on 6% urea 1 x TBE polyacrylamide gels at 60 W. After each extension of the sequence, new primers were designed and so on until the clones were completed. The sequence and orientation of the primers used is given in Appendix B.

The remaining part of the *crk1* gene was revealed by direct sequencing of the lambda clones (7.1 and 8.4) and using the Thermosequenase Radiolabeled Terminator Cycle Sequencing Kit (Amersham). The template concentration in the PCR labelling reaction was adjusted between 500 to 700 ng of lambda DNA which resulted in clear readings of the gels. The primer concentration was kept at 1 pmol for all reactions. The cycling termination reaction parameters were: 5 minutes initial denaturation at 95°C followed by 45 cycles of 95°C/30 sec, 55°C/ 1 min, 72°C/1.45 min. The reactions were terminated by the addition of stop solution (95% formamide, 20mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF). The samples were run on 6% urea glycerol tolerant polyacrylamide gels (UBS Amersham).

New primers were designed at the end of each new extension which allowed the sequencing of about 5.5 kb from each of the genomic DNA clones. The sequence of the primers used is shown in Appendix.

The two plasmid subclones (the respective Xba I fragments with sizes around 3.3 kb and 0.8 kb) could be partially aligned with a previously known cDNA fragment of the *crk1* (Kirby, 1996). According to the homology of the alignment, it was established that the two subclones were adjacent to each other on the genomic lambda phage clone (figure 4.6) with the Xba I restriction site being positioned in a intron between two exons of the *crk1* cDNA.

Further preliminary sequencing with the primers of the multiple cloning sites of other plasmid clones (e.g. the ~2.7 kb band) could not reveal more matches between the remaining known region of the cDNA and the genomic subclones. To facilitate the identification of the remainder of the *crk1* gene, the lambda genomic clones 7.1 and 8.4 were sequenced directly as described earlier in 4.2.1.4. Figure 4.6 clarifies the sequencing strategy and the orientation of the genomic sequence in relation to the Lambda FIX II vector arms.

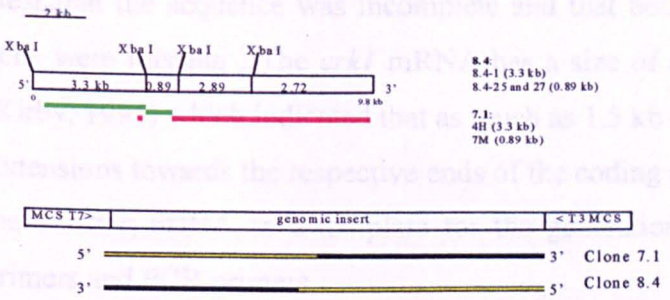


Figure 4.6 ***Crk1*** **sequencing and orientation of the genomic inserts in the Lambda FIX II vector arms.**

The two clones 7.1 and 8.4 carry the sequence of the *crk1* gene. About half of the total overall length of the sequenced DNA was covered by two plasmid subclones shown in the picture as green lines (4.2 kb). The rest of the *crk1* was revealed by sequencing the lambda DNA directly (red line). To prove that the two plasmid clones were following each other, a small fragment of the genomic clones in this region was resequenced directly (small red line). Relative to the MCS of the vector arms 7.1 is oriented with its 5'-end towards the T7 promoter whereas 8.4 is oriented with its 5'-end towards the T3 promoter within the MCS of vector.

The results from the sequencing of the plasmid subclones were combined with these from the lambda DNA sequencing to produce a fragment of about 9.8 kb for each of the two clones, 7.1 and 8.4.

During the course of sequencing, constant comparison was made between the newly uncovered genomic sequence and the 1.1 kb of partial cDNA of the *crk1* gene. The cDNA was used as a reference for "locating" the exons and introns on the genomic clones. The

cDNA was matched to a region on 5.59 kb on the genomic DNA for both lambda clones 7.1 and 8.4. The rest of the genomic sequence, 2.54 kb upstream of the 5.59 kb fragment, and 1.69 downstream of it, represented the novel part of the *crk1* gene. It was expected that the full genomic sequence was present on the 7.1 and 8.4 clones. To show this, it was necessary to discover where the beginning and end of the coding region lay. An attempt was made to determine the start and stop site using computer predictions but this approach gave uncertain results. It was decided to focus on the isolation of the 5'-and the 3'-end of the cDNA sequence of the *crk1*. The detailed procedure is presented in part 4.3 of the thesis.

4.3 Copy DNA isolation

The goal of this part of the project was to establish the complete cDNA sequence of the *crk1* gene. A 1.1 kb fragment of the *crk1* cDNA was determined (Kirby, 1996), however it was clear that the sequence was incomplete and that both ends of the coding region of the *crk1* gene were missing . The *crk1* mRNA has a size of approximately 2.6 kb on Northern blots (Kirby, 1996) which indicated that as much as 1.5 kb of the cDNA remained unknown.
Extensions towards the respective ends of the coding sequence of the *crk1* were planned using the existing cDNA as a template for the generation of gene specific reverse transcription primers and PCR primers.
This part of the thesis describes the isolation of the 5'-and the 3'-ends of the cDNA and the assembly of a full-length cDNA clone.

4.3.1 Isolation of the 5'-end of the *crk1* cDNA

A cRACE method (Maruyama *et al*, 1995) was adopted for the identification of the 5'-end of the *crk1* cDNA. In the cRACE (circular or concatameric first-strand cDNA-mediated RACE), the mRNA is reverse transcribed with a 5'-phosphorylated gene specific primer. The first strand cDNA is self-ligated to produce a circular molecule which is used in a PCR amplification to give the 5'-end of the mRNA. A diagram of the cRACE method is shown in figure 4.7 using as an example the first extension of the cDNA. In total, three consecutive extensions were carried out on the 5'-end of the *crk1* gene. The extensions are presented together as the only difference between them was in the cDNA synthesis primers and the primers for the PCR. Description of the procedure and the resulting PCR products are presented in part 4.3.5.

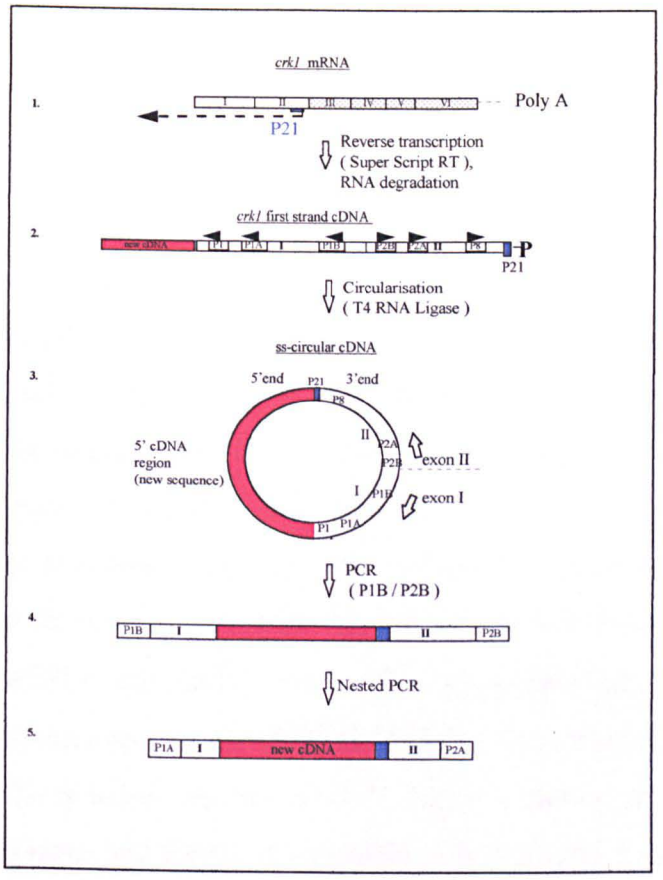


Figure 5.7. Isolation of the 5′ end of the *crk1* cDNA. Scheme of the cRACE method.

1) Reverse transcription of mRNA with 5′-phosphorylated gene specific primer P21.
2) 2) After hydrolysis of the mRNA template, the cDNA strand is ligated using T4 RNA ligase (3).
3) The expected 5′-end of the mRNA is given in red.
4) and 5) Two consecutive PCRs with gene specific primers: P1B and P2B, and P1A and P2A. The resulting PCR product is cloned into a vector and sequenced.

5.3.2 Messenger RNA isolation

mRNA was isolated using Dynabeads oligo dT_{25} (Dynal, UK). The maximum binding capacity of the beads is 2µg mRNA per extraction from 0.1g tissue. In order to make better use of the beads, 2.5 g of young seedlings were ground with the extraction buffer (100 mM Tris-HCl, pH 8.0, 500 mM LiCl, 10 mM EDTA, pH 8.0, 1% LiDS, 5 mM DTT), and homogenised with a pestle. Aliquots of 1 ml were pipetted into 1.5 ml Eppendorf tubes, which were quickly frozen in liquid Nitrogen. 200µl of Dynabeads were mixed with each 1 ml of homogenate and the mRNA was isolated following the instructions of the manufacturer. The same beads were reused four more times on new aliquots of the homogenate. The mRNA was separated into 200 µl aliquots and 3 volumes of 100% ethanol were added. The mRNA/ethanol mixture was stored at -20°C.

The mRNA was precipitated from the solution with 0.3 volumes of 3M sodium acetate and collected by centrifugation at 13 000 rpm for 15 minutes at 4°C. The pelleted RNA was washed with 70% ethanol and left to air dry for 1 minute. Then the mRNA was re-dissolved in 10 µl sterile distilled water and used in the reverse transcription reaction.

5.3.3 Reverse transcription

The SuperScript RT II reverse transcriptase (Gibco) was chosen for the reverse transcription of the mRNA following the instructions supplied with the enzyme.

Three consecutive reverse transcriptions of the 5'-end of the *crk1* mRNA were conducted using the following primers: p21-phosphorylated (part 1 of table 5.3), p1 new-phosphorylated (part 2 of table 5.3) and p5'-1cDNA (part 3 of table 5.3).

The mRNA (at a concentration of 1.5 to 2μg) and the 5' phosphorylated gene specific primer (table 5.3 and figures 5.8 and 5.9) were mixed in a total volume of 11 μl and heated to 70°C for 10 minutes. The tube was moved on ice, centrifuged briefly and the rest of the components of the reverse transcription reaction were added to the tube: 4μl 5x First Strand buffer, 2μl 0.1 M DTT, 1μl 10mM dNTPs. The components were mixed gently and incubated at 45°C for 2 minutes prior to the addition of 400 U (2μl) SuperScript RT II.

The reaction was left at 45°C for 50 minutes, the enzyme was inactivated at 70°C for 15 minutes and the mRNA template was degraded by addition of 2U RNase H and incubation at 37°C for 20 minutes.

The first strand cDNA was precipitated with 1 volume 4M ammonium acetate and 2 volumes iso-propanol and incubation at room temperature for 20 minutes. The cDNA was pelleted by centrifugation at 13 000 rpm for 10 minutes at room temperature, washed with 5 volumes 70% ethanol and centrifuged again for 10 minutes. The pellet was dried for 1 minute on air and re-dissolved directly in the ligation buffer.

5.3.4 Ligation of the single stranded cDNA

The first strand cDNA was dissolved in ten microlitres T4 RNA ligase mixture containing: 25% polyethylene glycol (mol. weight ~ 8000), 1 mM hexamine cobalt chloride, 0.01 mM ATP, 10mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 1 $\mu g ml^{-1}$ bovine serum albumin and 10 U T4 RNA ligase (New England BioLabs).

The ligation mixture was incubated at 22°C for 16 hours and an aliquot was taken directly for the PCR amplification of the 5'-end.

5.3.5 PCR amplification of the 5'-end of the *crk1* gene

A PCR reaction was set up containing 0.1μl of the T4 RNA ligase treated cDNA. The reaction was carried out for 35 cycles at 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute. The PCR product was diluted 10^{-3} fold and 2 μl of the dilution was used for the

nested PCR with the second set of gene specific primers. For all amplifications the concentration of the components in the PCR was kept constant excluding the amount of DNA template.

PCR reagents	
template DNA(single stranded DNA or PCR template) - table 5.3	1µl
25 mM $MgCl_2$	2 mM
10 x PCR buffer	1 x (5 µl)
10 mM dNTPs	200 µM
3′ Primer	100-150 pmol (1µl)
5′ Primer	100-150 pmol (1µl)
Taq DNA polymerase, (2.5 U μl^{-1})	1 U (0.4 µl)
H_2O	to final volume 50µl

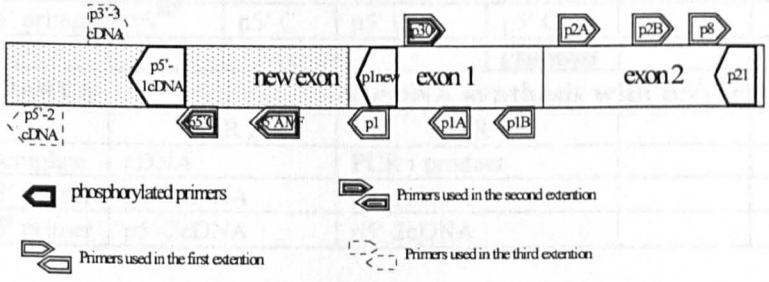


Figure 4.8

Scheme of the 5′-cDNA extension of the *crk1* gene. The new cDNA including the ATG start site has been determined using the cRACE technique. The new cDNA is shown in grey colour. The combination of primers used is also given in table 4.3.

In the first cDNA extension, several primer pairs were used for the amplifications PCRI (2 pairs) and PCR II (5 pairs). The second PCR produced a variety of bands with different sizes for each amplification reaction. Most of these bands were purified (Geneclean 2, Bio La Jolla), cloned into pGEM-T plasmid (Promega) and sequenced (Thermosequenase 2.0, UBS Amersham). It appeared that the difference in length of the PCR products was due to two reasons. The size of the 5′-end extension varied, and surprisingly, some of the sequenced inserts contained regions of the *crk1* cDNA downstream of the annealing site of the RT synthesis primer! This outcome is most likely due to unspecific annealing of the primer.

A second round of cDNA synthesis was carried out with phosphorylated primer P1-new (table 4.3, part 2) because it seemed that the first extension still did not reveal a potential ATG transcription start site. As a result a further 550 bp of the cDNA of the *crk1* gene was found and it appeared that this sequence contained a potential start plus 5′ untranslated leader sequence (figure 4.9). Further, a third cDNA extension (part 3, table 4.3) was conducted but

the product of this amplification seemed to share no homology with the genomic sequence of the *crk1* gene meaning that the cDNA sequence revealed in the first two extensions was indeed the entire 5′-end of the gene.

Table 4.3 **5′-extension of the *crk1* cDNA using the cRACE method.**

The three extensions are given as parts 1, 2 and 3 of this table

PART 1			**5′ cDNA synthesis with p21 (5′-phosphorylated)**				
	PCR I		PCR II				
template	cDNA		p2A-p1B PCR	p2A-p1B PCR	p2B-P1A PCR	p2A-p1B PCR	p2A-p1B PCR
3′ primer	p2A3′	p2B3′	p2B 3′	p8	p8	p2B 3′	p8
5′ primer	p1B5′	p1A5′	p1A 5′	p1	p1	p1B 5′	p1A 5′
							clone 5A
PART 2			**5′ cDNA synthesis with p1 new (5′-phosphorylated)**				
	PCR I		PCR II				
template	cDNA	cDNA	p30-pA^{MF} PCR	p30-p5′C PCR			
3′ primer	p30	p30	p30	p30			
5′ primer	pA^{MF}	p5′ C	p5′ C	p5′ C			
				clone 5B			
PART 3			**5′ cDNA synthesis with p5′-1cDNA (5′-phosphorylated)**				
	PCR I		PCR II				
template	cDNA		PCR I product				
3′ primer	p3′-3cDNA		p3′-3cDNA				
5′ primer	p5′-2cDNA		p5′-2cDNA				

The products from the PCR were ran on a 1.5% 1 x TAE agarose gel (figure 4.9) and the bands "genecleaned" from the agarose and cloned into the pGEM-T vector (Promega). The ligation reaction was carried out at 15°C for 2 hours followed by incubation at 75°C for 10 minutes in order to inactivate the ligase.

Ligation reaction	
reagents	
pGEM-T	50 ng
"gc" PCR band	100 - 200 ng
10 x Ligase buffer	1x
T4 DNA ligase	1.8 U
H_2O	to final volume 10 µl

Two µl from each ligation were used to transform 50 µl competent *E. coli* XL Blue II bacterial cells (Nishimura *et al.*). The colonies were selected on LB TAXI plates and 10 positives were taken for amplification in 10 ml liquid LB (Amp/Tet) overnight at 37°C. The plasmids were isolated with the Wizard mini prep kit (Promega), digested with Nco I + Sal I or Sac I + Sac II which cut in the multiple cloning site of the pGEM-T vector. The digests were run on 1% agarose gels and the clones containing inserts of the expected size were selected for sequencing.

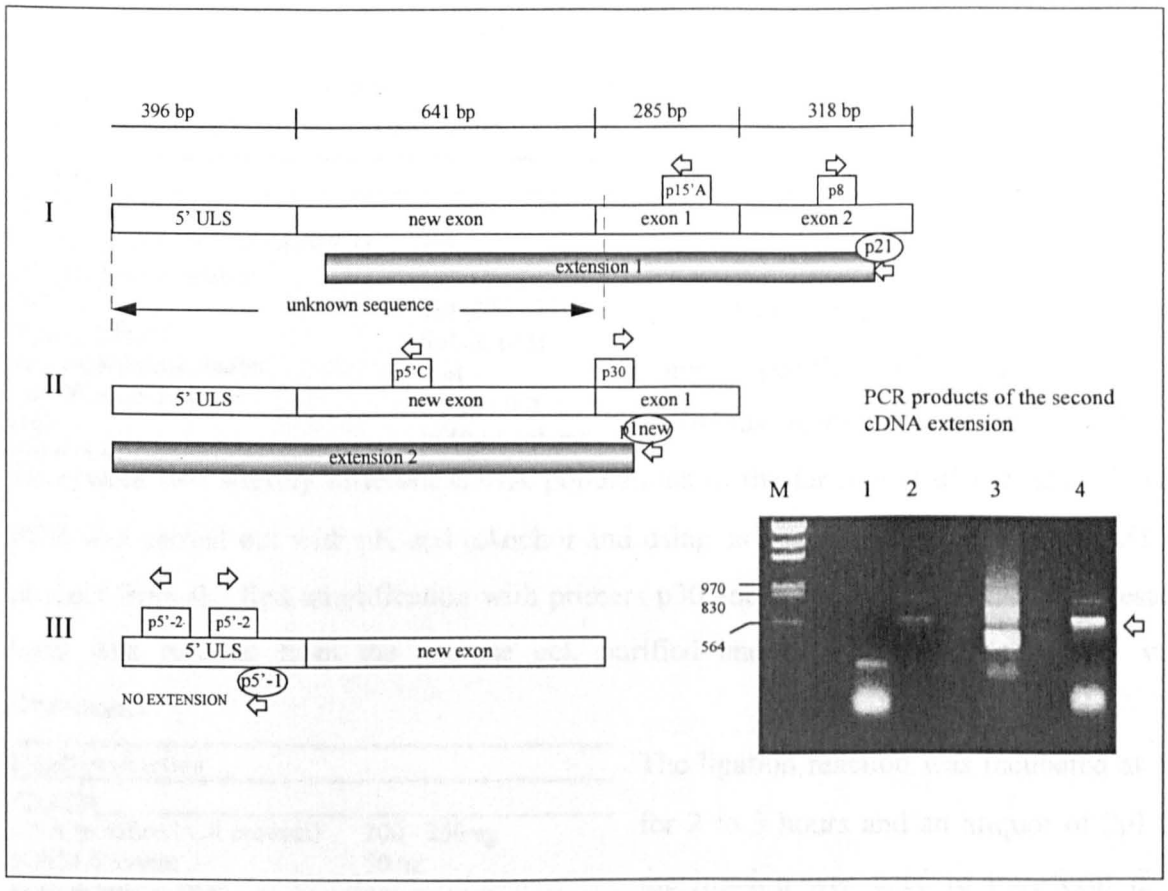


Figure 4.9 **5′ cRACE of *crk1* cDNA.**

Three consecutive extensions towards the 5′-end of the *crk1* mRNA were carried out resulting in the identification of the full-length 5′cDNA end.
The bars in the diagram represent the cDNA sequence, and exons shown are numbered as 1 and 2 before the identification of the "new" exon. The ovals underneath the sequence represent the primers used in the reverse transcription, and the rectangles above it represent the PCR primers. The agarose gel picture shown in the figure corresponds to the second extension: M- DNA size marker, 1 and 2 - first round of PCR amplification (p30-pAmf and p30-p5 respectively), 3 and 4 - second round of PCR (p30-p5′C for both reactions); PCR 1 is a template for the PCR in lane 3, and PCR in lane 2 is a template for the PCR in lane 4. 5′ULS – 5′ untranslated leader sequence.

4.3.6 Isolation of the 3′-end of the *crk1* cDNA

Using mRNA from young sugar beet seedlings as a template, first strand cDNA was synthesised with either oligo-dT_{36} or oligo(dT_{21})-Anchor primer in a reverse transcription reaction similar to the one described for the 5′-end of the *crk1* gene and using SuperScript RT II as before. The amount of mRNA used in the reaction was about 0.8 to 1μg and the primers were annealed to the mRNA at 70°C for 10 minutes. The reaction was incubated at 42°C for 50 minutes followed by inactivation of the enzyme at 70°C for 15 minutes.

Five μl of $1x10^{-3}$ dilution of the cDNA was taken directly for a PCR with one gene specific primer and the anchor primer. The PCR parameters were: initial denaturation 2 minutes at

94°C followed by 35 cycles of: 94°C for 20 seconds, 55°C annealing for 1 minute, and extension at 72°C for 1 minute. The contents of the PCR were:

PCR reagents	
cDNA (10^{-3} dilution)	5µl
3′ gene specific primer (table 4.3)	1µl
5′ gene specific primer	1µl
dNTPs	1µl (200 µM)
$MgCl_2$, 25 mM	3µl (3 mM)
10 x PCR thermo buffer	1 µl
Taq DNA polymerase	0.4 µl (2U)
H_2O	to 50 µl volume

The PCR products were run through 1% TAE agarose gel, purified with Geneclean Bio Kit, and sequenced by using 5µl of the total 10µl of the purified bands and several gene specific *crk1* primers. As it was obvious from the sequencing gels that there were two slightly different cDNA populations in the far 3′-end of the gene, a second PCR was carried out with pK and pAnchor and using as a template 1µl of the purified PCR product from the first amplification with primers p30 and pAnchor (table 4.4). The resulting band was isolated from the agarose gel, purified and subcloned into pGEM-T vector (Promega).

Ligation reaction reagents	
DNA (purified PCR product)	200 - 250 ng
pGEM-T vector	50 ng
ligation buffer, 10 x	1 x
T4 DNA ligase	3 U
H_2O	to final volume 10 µl

The ligation reaction was incubated at 16°C for 2 to 3 hours and an aliquot of 2µl from the ligation was used to transform *E. coli* XL1 Blue competent bacteria (Nishimura *et al.*). The positive colonies were selected on TAXI LB plates and propagated for plasmid isolation. The plasmids were isolated with Wizard miniprep kit (Promega). Following the plasmid isolation, a standard digest reaction of a microgram plasmid DNA with a combination of Nco I and Sal I enzymes in restriction buffer D (Promega) was carried out to verify the presence of inserts. The inserts were of two slightly different sizes. Five clones were selected for sequencing with Thermosequenase Version 2.0 sequencing kit (Amersham) and using the plasmid promoter primers T3 and SP6 as well as 5′ *crk1* gene specific primer pM. The sequences obtained were compared between themselves as well as with the genomic DNA of the *crk1* clones 7.1 and 8.4.

Several combinations of primers were used in the 3′ cDNA amplification. The primer combinations used are shown in table 4.4.

Table 4.4 **3′-end cDNA isolation.**

cDNA template	5′ primer	3′ primer	product
oligo(dT_{21})Anchor primer synthesised cDNA	p30	pAnchor	~ 1.8 kb
"	p8	pAnchor	~ 1.5 kb
"	pK	pAnchor	~ 0.6 kb
pAnchor-p30 PCR product	pK	pAnchor	~ 0.6 kb

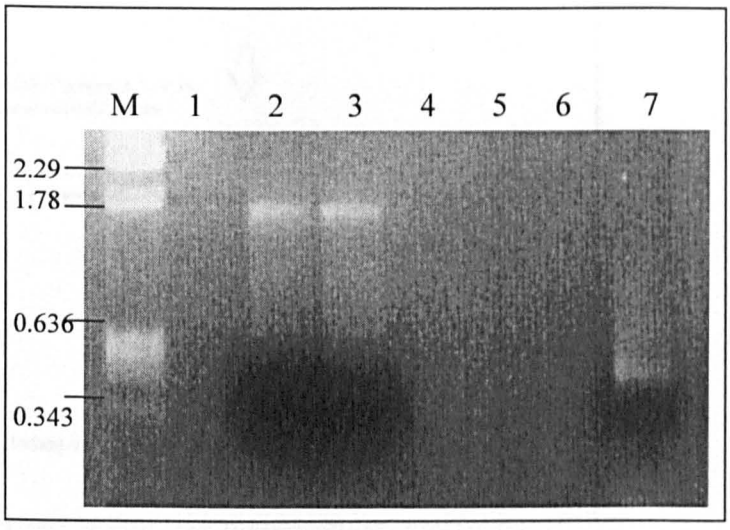


Figure 4.10. 3′ cDNA isolation.

M- DNA size marker, size of the marker shown in kb. Lane 2 and 3 - PCR amplification product on oligo(dT_{21}) synthesised cDNA with primer combination (p30 - pAnchor) and lane 7 - primers (pK and pAnchor). The products were excised from the gel and "genecleaned". They were sequenced directly, then the 0.6 kb product was reamplified and cloned into pGEM-T vector and sequenced.

Figure 4.11 presents an outline of the 3′end cDNA isolation procedure.

Some of the gene specific primers used for the PCR were included as a reference of the size of the resulting cDNA. In all PCR reactions, several controls were included. For the presence of the *crk1* message and a positive control for the PCR , a combination of internal *crk1* gene specific primers was used (p30-p23, p 5′crk1 -p3′ crk2) and their positions can be seen in the appendix giving the list and positions of the primers in the cDNA isolation. Also the usual negative controls containing no template, no primers or just one of the primers were included for all PCRs. Following the isolation of the 5′- and 3′- cDNA ends of the *crk1* a final PCR was carried out in order to amplify longer stretches of the *crk1* cDNAs. Several combinations of primers were used - they are shown in table 4.4 along with the sizes of the cDNAs produced. The PCRs were ligated into the MCS of the PCR cloning vector pGEM -T (Promega) exactly as the other PCR products. Some positive clones were selected for sequencing to verify the sequence composition.

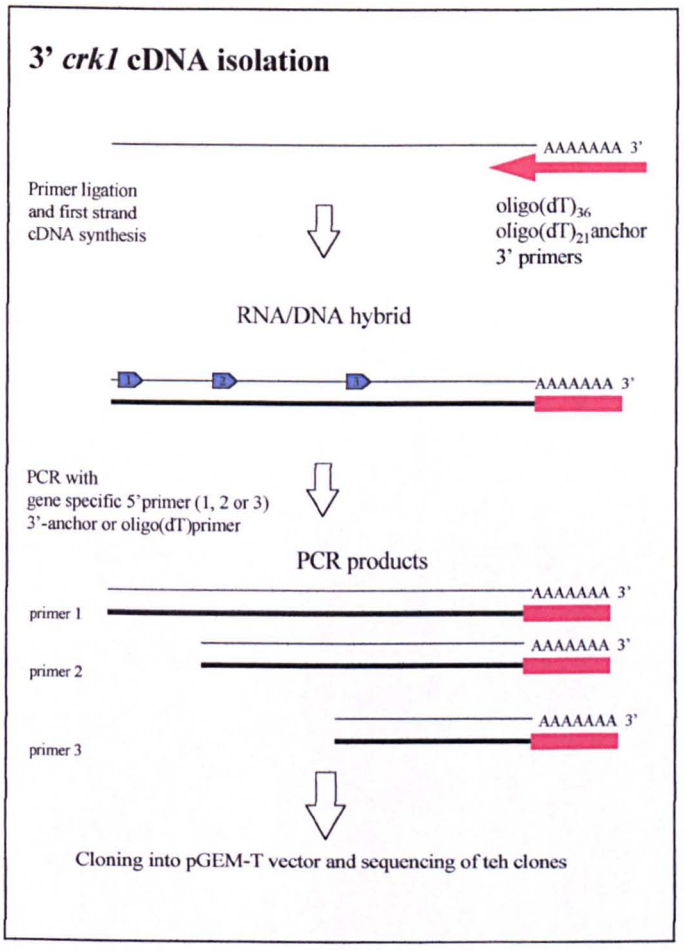


Figure 5.11. Isolation of the 3'-end of the *crk1* gene.

First strand cDNA was synthesised using oligo(dT_{36}) or oligo(dT_{21})Anchor primers for the reverse transcription. The RNA/DNA hybrid was taken for PCR with 5' gene specific primer (1, 2 or 3) and the Anchor primer. The products of the PCR were sequenced.

Table 5.5 PCR isolation of longer cDNA clones of the *crk1* gene.

template	5' primer	3' primer	product	plasmid clones
oligo(dT_{21})Anchor synthesised cDNA	p 2A 3'	Anchor	~1.3 kb	yes
oligo(dT_{36}) synthesised cDNA	p30	p23	~1.1 kb	no
oligo(dT_{36}) synthesised cDNA	p3'A ^{MF}	p23	~1.3 kb	yes
oligo(dT_{21})Anchor synthesised cDNA	p3'A ^{MF}	Anchor	~1.8 kb	yes

5.4 DNA sequence analysis

The programme packages GCG (Genetic Computer Group, Wisconsin) and FASTA (Bill Pearson's sequence similarity search programs) available on CLRS SEQNET, Daresbury

were used in the database search and retrieval (programmes LOOKUP, FETCH, STRINGSEARCH), fragment assembly (ASSEMBLE), sequence editing and comparison (SEQED, GAP, ALIGN) and sequence similarity search (FASTA). Aligned sequences were displayed with PRETTY, PRETTYBOX and PRETTYPLOT programs.

4.5 Results

4.5.1 Genomic DNA of the *Bvcrk1* gene

Two full-length genomic DNA clones of the *Bvcrk1* gene were isolated in the process of screening sugar beet Lambda FIX II genomic library with a PCR partial genomic sequence (Z70703) of the *crk1* gene (Kirby, 1996).

The two clones, designated 7.1 and 8.4, were identified out of six potential positives from the $3x10^5$ pfu screened. The presence of *crk1* homologous sequence was verified by southern blot analysis of the Xba I digested genomic clones. It was located on a 3.4-kb fragment, which was subcloned into pZErO-2 and sequenced (Sequenase 2.0, UBS, Amersham) from both directions of the multiple cloning site as well as with several internal *crk1* specific primers. Comparison between the 3.4 kb genomic subclone and the 1.1 kb partial *Bvcrk1* cDNA sequence (Kirby, 1996) showed that only a portion of the partial cDNA was present on the 3.4 kb Xba I fragment.

When the sequencing of the 3.4 kb Xba I subfragments of 7.1 and 8.4 was completed, it turned out that their 5'-ends contained the T3 (8.4) or T7 (7.1) promoter sites from the MCS of the Lambda FIX II vector. The unique Xba I site in the MCS of the arms lies after the T3/T7 promoter sites. This helped to work out that the genomic inserts were cloned in opposite direction. The 5'-end of 7.1 was positioned downstream from the left lambda arm, and the 5'-end of 8.4 was lying upstream from the right arm of the vector.

The size of the genomic inserts was calculated by adding up the sizes of the Xba I restriction fragments and was approximately 18.9 kb for 7.1 and 20.6 kb for 8.4.

The first plasmid subclone (the 3.4 kb subfragment) was obviously not producing any new information with regard to the coding sequence of the *crk1* gene. Further sequencing had to be carried out on the remaining plasmid subclones of 7.1 and 8.4 in an attempt to find the next fragment of DNA with homology to the remainder of the cDNA sequence. As a start, the Xba I subclones in pZErO-2 were sequenced with M13 forward and M13 reverse primers from the multiple cloning site of the plasmid vector. The sequence information was aligned with the piece of *crk1* cDNA and luckily, a small 0.89-kb subfragment matched in part with the

remaining cDNA. This second fragment seemed to follow the 3.3 kb one. The 3'-end of the 3.3 kb subfragment ended in an intron region of the *crk1* DNA and the 5'-end of the 0.89 kb began in the intron sequence, which made it impossible to be certain that both fragments were indeed following each other on the genomic clones. The easiest way to check this was to try and sequence the genomic lambda clones directly using suitable gene specific primers. The clones were taken as templates in a sequencing reaction (Thermosequenase 2.0, UBS Amersham) using a modification of the protocol advised by the suppliers. The approach seemed to work very well and it verified the expectation that the 3.4 kb subfragment was immediately followed by the 0.89 kb on both 7.1 and 8.4. Having been successful with the new sequencing strategy, the further identification and sequencing of plasmid subclones was dropped and replaced by the "direct" sequencing of the lambda genomic clones. In this fashion, using a succession of *crk1* specific primers approximately another 5.5 kb of the genomic sequences of 7.1 and 8.4 clones was obtained.

The total length of sequenced genomic DNA from the lambda clones was 9.75 kb and 9.83 kb for 7.1 and 8.4 respectively. The clones were aligned using the GCG programme GAP (figure 4.13). The overall identity of the two DNAs was 97.5% with mismatches and gaps occurring mainly in the introns of the gene.

A previously preformed Southern blot hybridisation analysis of the *Bvcrk1* gene is presented in figure 4.12 (M. J. Kirby, Ph.D. thesis 1995). In the triploid sugar beet variety, which was also used for the construction of the genomic library, it is clear that the *Bvcrk1* gene is present at a low copy number. This analysis supports the results obtained in this study with the isolation of two neraly identical clones for the *Bvcrk1* from the Lambda FIX II genomic library.

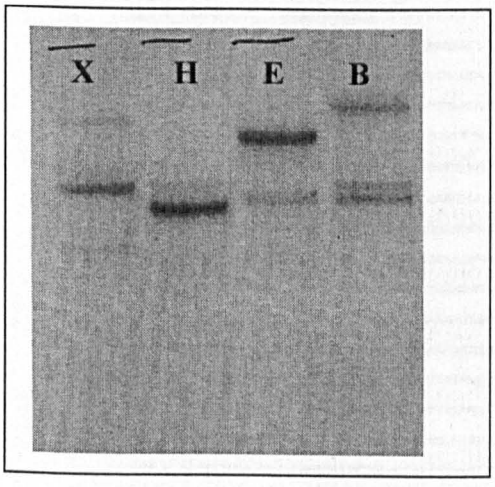


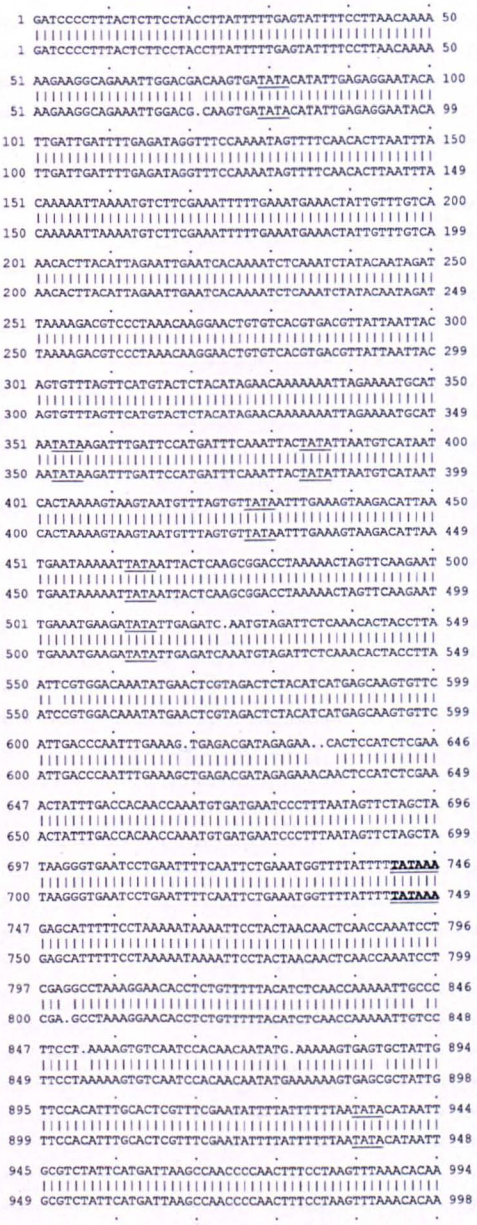
Figure 4. 12 *Bvcrk1* probed Southern blot of genomic DNA from sugar beet, cultivar Regina, a triploid variery.

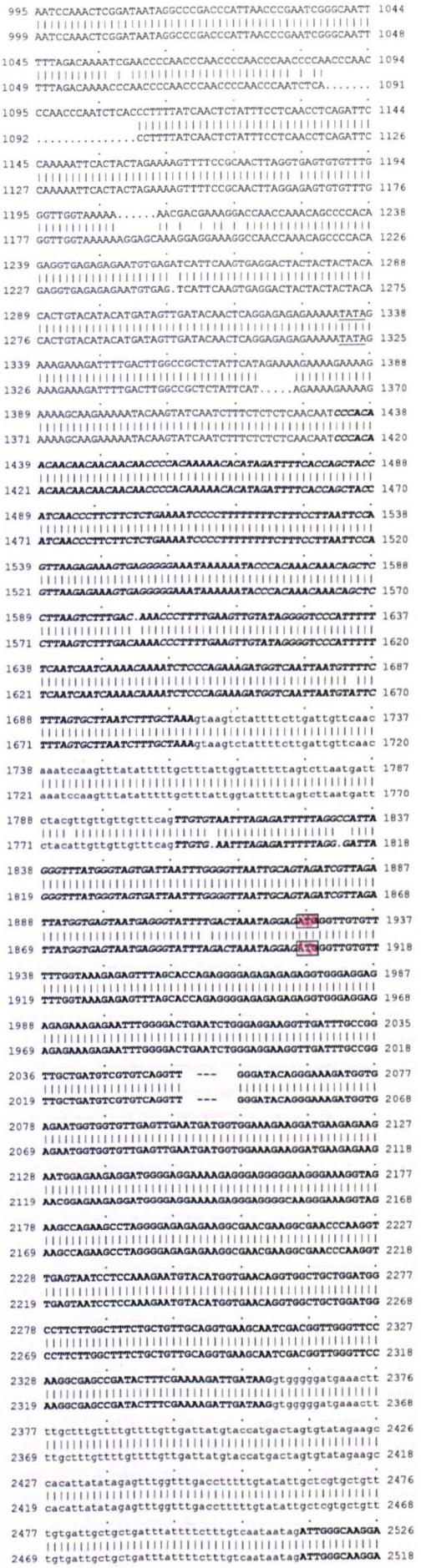
10µg of genomic DNA digested with one of the following restriction enzymes before agarose gel separation and Southern hybridisation: Xba I (X), Hind III (H), EcoR I (E), BamH I (B).

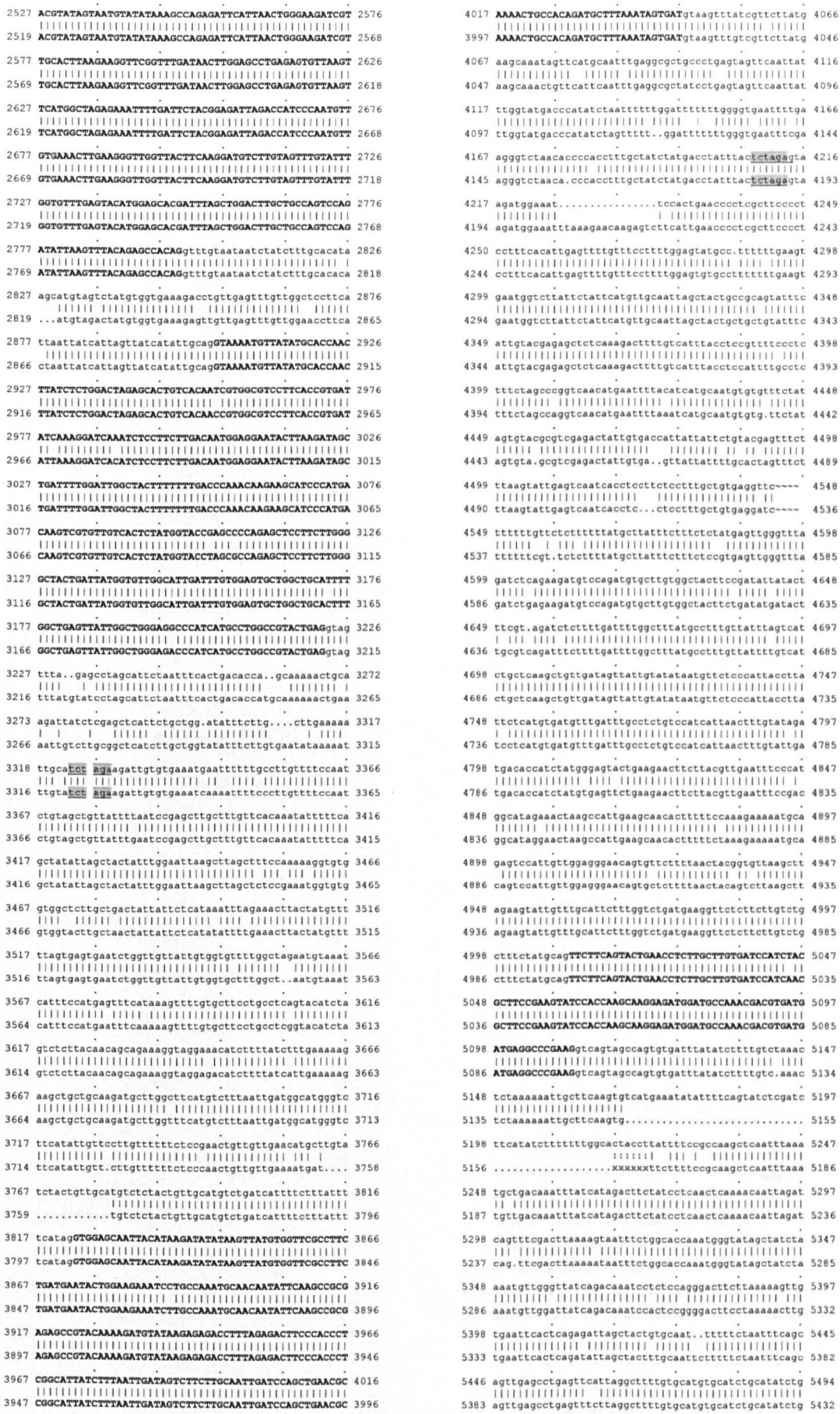
Picture reprinted from M. J. Kirby, Ph.D. thesis, 1995.

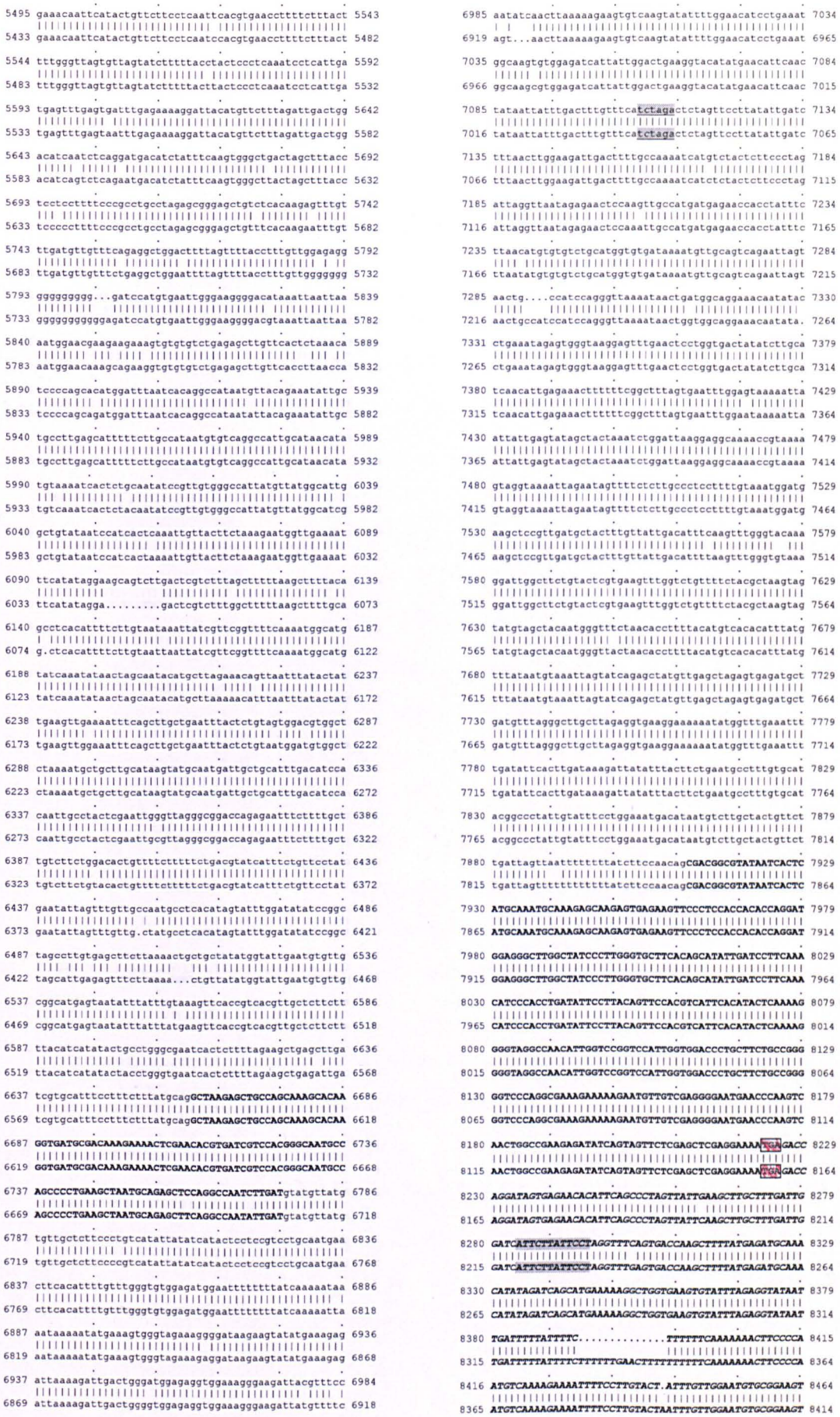
Figure 4.13 Alignment of the two B*vcrk1* sequences from the genomic clones - 7.1 and 8.4 using the GCG programme GAP.

The upper nucleotide sequence in the alignment represents 8.4, The lower – 7.1.
The cDNA region is shown with bold upper case straight letters and the intron regions are given in lower case. The ATG start codon and the stop sites are boxed. The region showed in upper case italics underlines the non-coding part of the cDNA sequence. The TATA box consensus sequence TATAa/tAa/T is double underlined and in bold. Other TATA sequences in the promoter region of the gene are underlined. The Xba I restriction sites are underlined and highlighted in grey. The gap introduced in both strands in the second column of this page does not represent any brake of the sequence and was inserted in order to avoid realigning the sequences manually.









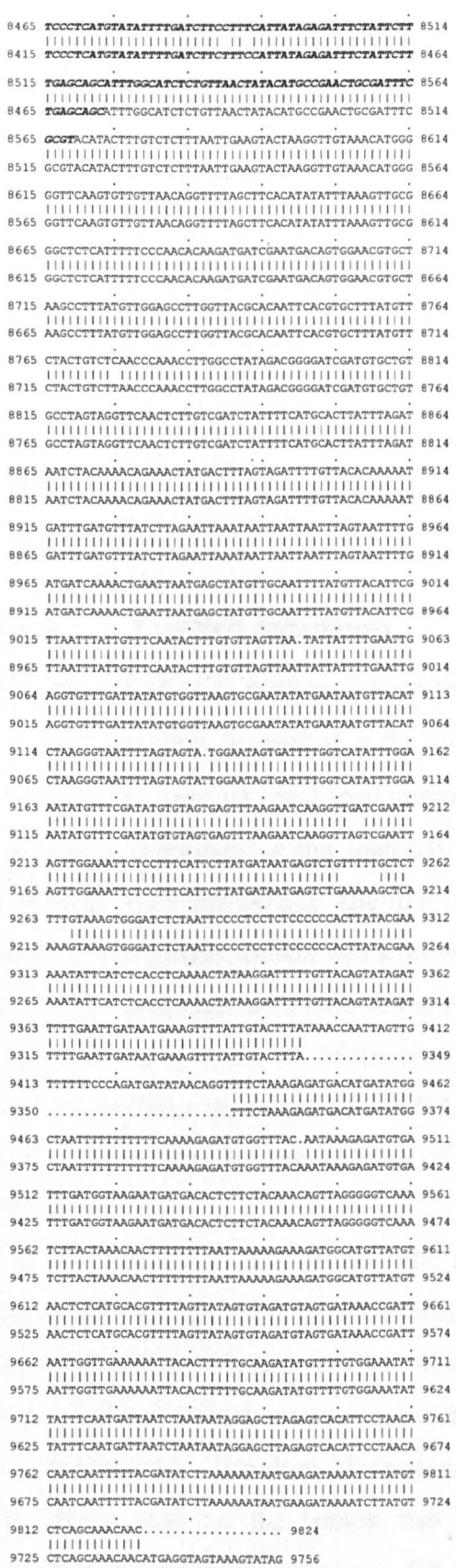


Figure 4.13 Alignment of the two genomic clones of the *Bvcrk1* gene 7.1 and 8.4.
(End of figure)

Comparison between the 1.1 kb partial cDNA sequence of *crk1* and the genomic data made it easy to determine the position of 6 exons. Clearly, the partial cDNA was lacking the 5′-end of its coding region because it was starting in the ATP-binding site region of the kinase domain. It also seemed to be short of its 3′ -end as no stop codon seemed to be present. It was unclear how long the missing parts on both ends of the cDNA might be. The prediction of the possible positions of the ATG start sites at the 5′-end of the genomic clones (the 3.3 Xba I fragments) did not prove very conclusive, neither did those aimed at determining where the 3′-end lay as the predictions resulted in multiple possibilities for both start and stop sites. Thus, full-length cDNA had to be isolated, or at least its 5′- and 3′-ends so as to be able to reconstruct a full-length cDNA of the *crk1* gene.

4.5.2 5′-cDNA extension

The elegant cRACE method (Maruyama *et al.*, 1995) was used to obtain the missing 5′-end of the gene. It resulted in the identification of the first and so far unknown exon of the *Bvcrk1* gene and its 5′-end untranslated leader sequence.

As it was described in the methods section of this chapter, three consecutive 5′-extensions were carried out. The first two revealed new 835 bp of the *crk1* cDNA. The result of the third extension was a non-specific PCR product.

It was accepted that the 5′-end of the cDNA was reached with the second 5′-extension (figure 4.14 and 4.15). All of the clones from the second extension ended at the same position (relative to the genomic DNA used as reference) with an added G at their 5′-end, characteristic for the capping of the eukaryotic mRNA (Weaver and Hendrick, 1991).

The fact that the third 5′-cRACE did not result in gene specific sequence is possibly due to non-specific PCR amplification. The PCR product appeared on the agarose gel as a smeared band between 400 to 450 bp in size. The cloning of the product resulted in the isolation of only one recombinant bacterial clone. When sequenced, it proved to be non-specific. The low cloning efficiency together with the sub-optimal PCR conditions may be the reason that the last 5′-extension of the *crk1* gene was unsuccessful in terms that the it did not reproduce the 125 bp fragment which was expected if the experiment had worked well.

An alignment of five of the cloned 5′ cRACE fragments from the second extension is presented in figure 4.14. Two of the five clones shown in the alignment have a base pair substitution at position 496 where A is substituted with T. This introduces TAA stop site in the coding sequence of the gene. The remaining three sequences from the alignment are homologous to the genomic sequence of the *crk1* gene in this region. The last nucleotide of the 5′ -extension is given in bold and underlined. The sequence upstream the 5′-end of the extension is in fact the reverse transcription phosphorylated primer used for the synthesis of the cDNA.

Figure 4.14 Alignment of the 5′-ends of the *Bvcrk1* cRACE products

The end of the first extension is shown next to the sequences in the alignment and is in bold (A at 500 bp). Nucleotide differences in the sequences are highlighted. The ATG start codon is boxed as well as the TAA stop site present in two of the sequences.

```
        1                                                                   70
cDNA3   ~~~TGTATATAAA GCCGAGATTC ATTAACTGGG AGCCCAC AACAACAACA ACAACCCACA CAAAAACACA   G - Start
cDNA4   ~~~~GTATATAAA GCCGAGATTC ATTAACTGGG AGCCCAC AACAACAACA ACAACCCCCA CAAAAACACA   of the 5'cDNA
cDNA7   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~GCCCAC AACAACAACA ACAACCCCCA CAAAAACACA   clones
cDNA6   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
cDNA5   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        71                                                                  140
cDNA3   TAGATTTTCA CCAGCTACCA ACAACCCTTC TTCTCTGAAA ATCCCCTTTT TTTTCTTTCC TTAATTCCAG
cDNA4   TAGATTTTCA CCAGCTACCA TCAACCCTTC TTCTCTGAAA ATCCCCTTTT TTTTCTTTCC TTAATTCCAG
cDNA7   TAGATTTTCA CCAGCTACCA TCAACCCTTC TTCTCTGAAA ATCCCCTTTT TTTTCTTTCC TTAATTCCAG
cDNA6   ~AGATTTTCA CCAGCTACCA TCAACCCTTC TTCTCTGAAA ATCCCCTTTT TTTTCTTTCC TTAATTCCAG
cDNA5   ~AGATTTTCA CCAGCTACCA TCAACCCTTC TTCTCTGAAA ATCCCC.TTT TTTTCTTTCC TTAATTCCAG

        141                                                                 210
cDNA3   TTAAGAGAAA GTGAGGGGGA AATAAAAAAT ACCCACAAAC AAACAGCTCC TTAAGTCTTT GACAAACCCT
cDNA4   TTAAGAGAAA GTGAGGGGGA AATAAAAAAT ACCCACAAAC AAACAGCTCC TTAAGTCTTT GACAAACCCT
cDNA7   TTAAGAGAAA GTGAGGGGGA AATAAAAAAT ACCCACAAAC AAACAGCTCC TTAAGTCTTT GACAAACCCT
cDNA6   TTAAGAGAAA GTGAGGGGGA AATAAAAAAT ACCCACAAAC AAACAGCTCC TTAAGTCTTT GACAAACCCT
cDNA5   TTAAGAGAAA GTGAGGGGGA AATAAAAAAT ACCCACAAAC AAACAGCTCC TTAAGTCTTT GACAAACCCT

        211                                                                 280
cDNA3   TTTGAAAGTT GTATAGGGGT CCCATTTTTT CAATCAATCA AACAAAATCT CCCAGGAAGA TGGTCAATTA
cDNA4   TTTGAAAGTT GTATAGGGGT CCCATTTTTT CAATCAATCA AACAAAATCT CCCAGAAAGA TGGTCAATTA
cDNA7   TTTGAAAATT GTATAGGGGT CCCATTTTTT CAATCAATCA AACAAAATCT CCCAGAAAGA TGGTCAATTA
cDNA6   TTTGAAAATT GTATAGGGGT CCCATTTTTT CAATCAATCA AACAAAATCT CCCAGAAAGA TGGTCAATTA
cDNA5   TTTGAAAGTT GTATAGGGGT CCCATTTTTT CAATCAATCA AACGAAATCT CCCAGAAAGA TGGTCAATTA

        281                                                                 350
cDNA3   ATGTTTTCTT TAGTGCTTAA TCTTTGCTAA ATTGTGAATT TAGAGATTTT TAGGGATTAG GGTTTATGGG
cDNA4   ATGTTTTCTT TAGTGCTTAA TCTTTGCTAA ATTGTGAATT TAGAGATTTT TAGGGATTAG GGTTTATGGG
cDNA7   ATGTTTTCTT TAGTGGTTAA TCTTTGCTAA ATTGTGAATT TAGAGATTTT TAGGGATTAG GGTTTATGGG
cDNA6   ATGTTTTCTT TAGTGCTTAA TCTTTGCTAA ATTGTGAATT TAGAGATTTT TAGGGATTAG GGTTTATGGG
cDNA5   ATGTTTTCTT TAGTGCTTAA TCTTTGCTAA ATTGTGAATT TAGAGATTTT TAGGGATTAG GGTTTATGGG

        351                                                                 420
cDNA3   TAGTGATTAA TTTGGGGTTA ATTGCAGTAG ATCGTTAGAT TATGGTGAGT AATGAGGGTA TTTAGACTAA
cDNA4   TAGTGATTAA TTTGGGGTTA ATCGCAGTAG ATCGTTAGAT TATGGTGAGT AATGAGGGTA TTTAGACTAA
cDNA7   TAGTGATTAA TTTGGGGTTA ATTGCAGTAG ATCGTTAGAT TATGGTGAGT AATGAGGGTA TTTAGACTAA
cDNA6   TAGTGATTAA TTTGGGGTTA ATTGCAGTAG ATCGTTAGAT TATGGTGAGT AATGAGGGTA TTTAGACTAA
cDNA5   TAGTGATTAA TTTGGGGTTA ATTGCAGTAG ATCGTTAGAT TATGGTGAGT AATGAGGGTA TTTAGACTAA

        421     M                                                           490
cDNA3   ATAGGAGATG GGTTGTGTTT TTGGTAAAGA GAGTTTAGCA CCAGAGGGGA GAGAGAGAGG TGGGAGGAGA
cDNA4   ATAGGAGATG GGTTGTGTTT TTGGTAAAGA GAGTTTAGCA CCAGAGGGGA GAGAGAGAGG TGGGAGGAGA
cDNA7   ATAGGAGATG GGTTGTGTTT TTGGTAAAGA GAGTTTAGCA CCAGAGGGGA GAGAGAGAGG TGGGAGGAGA
cDNA6   ATAGGAGATG GGTTGTGTTT TTGGTAAAGA GAGTTTAGCA CCAGAGGGGA GAGAGAGAGG TGGGAGGAGA
cDNA5   ATAGGAGATG GGTTGTGTTT TTGGTAAAGA GAGTTTAGCA CCAGAGGGGA GAGAGAGAGG TGGGAGGAGA

        491       end of First 5'extension                                  560
cDNA3   GAGAAAGAGA ATTTGGGGAC TGAATCTGGG AGGAAGG~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~   overlap with
cDNA4   GAGAAAGAGA ATTTGGGGAC TGAATCTGGG AGGAAGG~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~   the first
cDNA7   GAGTAAGAGA ATTTGGGGAC TGAATCTGGG AGGAAGGTTG ATTTGCCGGT TGCTGATGTC GTGTCAGGTT   5' cDNA
cDNA6   GAGTAAGAGA ATTTGGGGAC TGAATCTGGG AGGAAGGTT~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~   extension
cDNA5   GAGAAAGAGA ATTTGGGGAC TGAATCTGGG AGGAAGGTT~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        561                                 593
cDNA3   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~
cDNA4   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~
cDNA7   GGGATACAGG GAAAGATGGT GAGAATGGTG GTG
cDNA6   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~
cDNA5   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~
```

4.5.3 3′-cDNA extension

In parallel with the 5′-end extension, 3′-end extension of the *Bvcrk1* cDNA was carried out. Five 3′-end cDNA sequences from the 3′-end PCR (primers pair pK and pAnchor)

are presented below. There are two species of DNA sequence in the 3′-end of the *crk1* gene. They match the difference between the genomic clones of the *crk1* in the 3′-untranslated region of the gene. There is a 14 bp deletion in the 7.1 clone in comparison with 8.4, and it is also present in two of the five 3′-end extensions shown in the alignment in figure 4.15. Variability of the length of the 3′-end was observed and putative polyadenylation signal sequences were not detected in the expected vicinity of the poly (A) tail addition.

Figure 4.15 Alignment of the 3′-end extensions of the *crk1* gene.

The sequences presented are products of PCR using gene specific primer pK and pAnchor primer for the amplification of the 3′-end of the cDNA. The differences from the consensus are boxed in gray colour. The first two sequences in the alignment: 3end5 and 3end4 are homologous to the genomic sequence of the *crk1* clone 7.1. The remaining three sequences from the alignment are matching the other genomic clone - 8.4.

```
      1                                                                    80
3end5 ACGTCATTCA CATACTCAAA AGGGGTAGGC CAACATTGGT CCGGTCCATT GGTGGACCCT GCTTCTGCCG GGGGTCCCAG
3end4 ACGTCATTCA CATACTCAAA AGGGGTAGGC CAACATTGGT CCGGTCCATT GGTGGACCCT GCTTCTGCCG GGGGTCCCAG
3end3 ACGTCATTCA CATACTCAAA AGGGGTAGGC CAACATTGGT CCGGTCCATT GGTGGACCCT GCTTCTGCCG GGGGTCCCAG
3end2 ACGTCATTCA CATACTCAAA AGGGGTAGGC CAACATTGGT CCGGTCCATT GGTGGACCCT GCTTCTGCCG GGGGTCCCAG
3end1 ACGTCATTCA CATACTCAAA AGGGGTAGGC CAACATTGGT CCGGTCCATT GGTGGACCCT GCTTCTGCCG GGGGTCCCAG

      81                                                                  160
3end5 GCGAAAGAAA AAGAATGTTG TCGAGGGGAA TGAACCCAAG TCAACTGGCC GAAGAGATAT CAGTAGTTCT CGAGCTCGAG
3end4 GCGAAAGAAA AAGAATGTTG TCGAGGGGAA TGAACCCAAG TCAACTGGCC GAAGAGATAT CAGTAGTTCT CGAGCTCGAG
3end3 GCGAAAGAAA AAGAATGTTG TCGAGGGGAA TGAACCCAAG TCAACTGGCC GAAGAGATAT CAGTAGTTCT CGAGCTCGAG
3end2 GCGAAAGAAA AAGAATGTTG TCGAGGGGAA TGAACCCAAG TCAACTGGCC GAAGAGATAT CAGTAGTTCT CGAGCTCGAG
3end1 GCGAAAGAAA AAGAATGTTG TCGAGGGGAA TGAACCCAAG TCAACTGGCC GAAGAGATAT CAGTAGTTCT CGAGCTCGAG

      161                                                                 240
3end5 GAAAATGAGA CCAGGATAGT GAGAACACAT TCAGCCCTAG TTATTGAAGC TTGCTTTGAT TGGATCATTC TTATTCCTAG
3end4 GAAAATGAGA CCAGGATAGT GAGAACACAT TCAGCCCTAG TTATTGAAGC TTGCTTTGAT TGGATCATTC TTATTCCTAG
3end3 GAAAATGAGA CCAGGATAGT GAGAACACAT TCAGCCCTAG TTATTGAAGC TTGCTTTGAT TGGATCATTC TTATTCCTAG
3end2 GAAAATGAGA CCAGGATAGT GAGAACACAT TCAGCCCTAG TTATTGAAGC TTGCTTTGAT TGGATCATTC TTATTCCTAG
3end1 GAAAATGAGA CCAGGATAGT GAGAACACAT TCAGCCCTAG TTATTGAAGC TTGCTTTGAT TGGATCATTC TTATTCCTAG

      241                                                                 320
3end5 GTTTCAGTGA CCAAGCTTTT ATGAGATGCA AACATATAGA TCAGCATGAA AAAGGCGGGT GAAGTGTATT TAGAGGTATA
3end4 GTTTCAGTGA CCAAGCTTTT ATGAGATGCA AACATATAGA TCAGCATGAA AAAGGCTGGT GAAGTGTATT TAGAGGTATA
3end3 GTTTCAGTGA CCAAGCTTTT ATGAGATGCA AACATATAGA TCAGCATGAA AAAGGCTGGT GAAGTGTATT TAGAGGTATA
3end2 GTTTCAGTGA CCAAGCTTTT ATGAGATGCA AACATATAGA TCAGCATGAA AAAGGCTGGT GAAGTGTATT TAGAGGTATA
3end1 GTTTCAGTGA CCAGGCTTTT ATGAGATGCA AACATATAGA TCAGCATGAA AAAGGCTGGA GAAGTGTATT TAGAGGTATA

      321                                                                 400
3end5 ATTGATTTTT ATTTTC.... .......... TTTTTTCAAA AAAACTTCCC CAATGTCAAA AGAAAATTTT CCTTGTACTA
3end4 ATTGATTTTT ATTTTC.... .......... TTTTTTCAAA AAAACTTCCC CAATGTCAAA AGAAAATTTT CCTTGTACTA
3end3 ATTGATTTTT ATTTTCTTTT TTGAACTTTT TTTTTTCAAA AAAACTTCCC CAATGTCAAA AGAAAATTTT CCTTGTACTA
3end2 ATTGATTTTT ATTCTC.TTT TTGAACTTTT TTTTTTCAAA AAAACTTCCC CAATGTCAAA AGAAAATTTT CCTTGTACTA
3end1 ATTGATTTTT ATTTTCTTTT TTGAAC.TTT TTTTTTCAAA AAAACTTCCC CAATGTCAAA AGAAAATTTT CCTTGTACTA

      401                                                                 480
3end5 TTTGTTGGAA TGTGCGGAAG TTCCCTCATG T~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
3end4 TTTGTTGGAA TGTGCGGAAG TTCCCTCATG TATATTTTGA TCTTCCTTTC ATTATAGAGA TTTCTATTCT TTGAGCAGCA
3end3 TTTGTTGGAA TGTGCGGAAG TTCCCTCATG TATATTTTGA TCTTCCTTTC ATTATAGAGA TTTCTATTCT TGGAGCAGC~
3end2 TTTGTTGGAA TGTGCGGAAG TTCCCTCATG TATATTTTGA TCTTCCTTTC ATT~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
3end1 TTTGTTGGAA TGTGCGGAAG TTCCCTCATG TATATTTTGA TCTTCCTTTC ATTATAGAGA TTTCTATTCT TGGAGCAGC~

      481                                  525
3end4 TTTGGCATCT CTGTTAACTA TACATGCCGA ACTGCGATTT CGCGT
```

4.5.4 Full-length cDNA of the *crk1* gene

Isolation of full-length cDNA from mRNA was not attempted; instead a full-length cDNA was reconstructed from two overlapping PCR partial clones. The larger portion in the assembly (1.8 kb) came from the sequence of the PCR products (amplification with primers $pA3'^{MF}$ and pAnchor on mRNA template), and the smaller part at the 5′-end was taken from the sequence of the 5′ cRACE extension products.

The two species of cDNAs match the two *crk1* genomic DNA sequences. The cDNA of the *Bvcrk1* gene corresponding to the genomic clone 8.4 is 2503 bp long, and the one matching 7.1 is 2535 bp. The identified clones seem to be two alleles of the same sugar beet gene. Apart from a 14 bp deletion at the 3′-untranslated region between the two cDNAs, only several conservative nucleotide substitutions are found in the coding region of the gene. These differences are presented in table 4.5 below. A slight difference occurs in the 3′-untranslated regions because of their varying length and several nucleotide substitutions and single base pair gaps.

Table 4.5 **Base pair substitutions in the coding sequence of the *crk1* gene.**

exon		codon change	amino acid
		8.4 / 7.1	8.4 / 7.1
third exon	1	AAT→AAC	N
	2	ATC→ATT	I
	3	GCC→GCG	A
	4	AGG→AGA	R
fourth exon	5	CTC→CTT	L
fifth exon	6	TCT→TCA	S
sixth exon	7	CTC→CTT	L
	8	CTT→ATT	L→I

A feature of the *Bvcrk1* genomic DNA is the presence of a 343 bp 5′-untranslated leader sequence and the lack of a TATA box in proximity to the transcription start site. Nevertheless, several potential TATA consensus sequences were located further upstream on the remaining 1.4 kb of genomic sequence. The closest best fit for TATA box matching the consensus TATAa/tAa/t (Wu *et al.*, 1995) is positioned at 740-747 bp (8.4) and 743-750 bp (7.1). The position of the transcription start site was determined at 1432 bp (8.4) and 1415 bp (7.1) where the second round cRACE extensions ended with an additional G nucleotide at their very 5'-end. This extra nucleotide is not present on the genomic DNA but was found on all of the cRACE

extensions. Its presence at the 5'-end of the mRNAs matches the position of the post-transcriptionally added 7-methylguanosine residue in the capping of the mRNA. This site seems to be correct also because of the presence of conserved eukaryotic transcription start site sequence CTCAACA (Hughes, 1996) lying 3 to 9 bp upstream from the defined position of the start.

Both cDNA clones contain open reading frame of 1797 bp predicted to encode a protein of 599 amino acids. The gene consists of seven coding exons separated by eight introns. A ninth intron was found in the 5'-untranslated leader sequence which consists of an entire untranslated exon plus the untranslated part of the first coding exon. The 5'- and 3'-splice junctions between the exons and introns followed the conserved dinucleotide consensus sequence GT and AG respectively (Brown, 1986).

The full-length cDNA of the *Bvcrk1* gene shown on figure 4.16 is the sequence matching the genomic clone 8.4. Its predicted protein sequence was used in the analysis of the primary structure and in the homology searches for the CRK1 protein (4.5.5). Figure 4.12 shows the alignment of the 7.1 and 8.4 with the exons presented in upper case and the introns in lower case.

Interestingly, one of the closest homologues of the *Bvcrk1* gene (overall identity of the predicted proteins of 63.5%), a predicted *Arabidopsis* cdk-like gene At1g54610, has a deduced amino acid sequence of 572 amino acids. The position of the exons, their length and the positions of the intron splice sites between the two gene neatly match, although the size of the introns is proportionally bigger in the sugar beet gene (see table 4.7). There are a few homologous plant sequences with a similar structure to *crk1*. An alignment of the predicted protein sequences from the *Arabidopsis* genomic sequences and several cDNAs from rice are presented in the following section of this chapter.

Table 4.6 Intron splice junctions in the two crk1 genomic clones

The 5′- and 3′-splice junctions are denoted with (:). The last codon (5′) and the first codon (3′) in each exon are shown in capital letters. The intron, shown as its first five nucleotides (5′) and its last five nucleotides (3′) is given in small letters. The figures bellow each intron show its position in the 7.1 and 8.4 cloned and its size in bp.

intron		5′-splice junction	Intron	3′- splice junction
		AAA:**gt**aag		ttc**ag**:TTG
1	7.1	1692....	97 bp	1789
5′-	8.4	1711....	97 bp	1808
UTR				
		AAG:**gt**ggg		aat**ag**:GTA
2	7.1	2345....	153 bp	2498
	8.4	2364....	153 bp	2517
		CAG:**gt**agt		tgc**ag**:GTA
3	7.1	2784....	101 bp	2885
	8.4	2803....	104 bp	2907
		CAG:**gt**aag		cat**ag**:GTG
4	7.1	3204....	589 bp	3793
	8.4	3226....	599 bp	3825
		GAT:**gt**aag		tcg**ag**:TTC
5	7.1	4023....	964 bp	4987
	8.4	4055....	956 bp	5011
		AAG:**gt**cag		tcg**ag**:GCT
6	7.1	5089....	1486 bp	6575
	8.4	5113....	1551 bp	6664
		GAT:**gt**atg		aac**ag**:CGA
7	7.1	6700....	1136 bp	7836
	8.4	6780....	1134 bp	7914

exon	Bvcrk1	At1g54610
1	435 (145)	363 (121)
2	288 (96)	285 (95)
3	319 (106)	318 (106)
4	227 (76)	222 (74)
5	102 (34)	95 (32)
6	114 (38)	110 (37)
7	311 (104)	320 (107)

Table 4.7 Comparison between the *Bvcrk1* and At1g54610

The first figure is the size of the corresponding exons in bp, and in brackets is the size of the deduced amino acid sequence.
The second, third and fourth exons contain the kinase core domain (subdomains I to XI). The first exon carries the large N terminal portion of the protein with a possible NLS. A second NLS is located in the last exon. The N-terminal domain as well as the protein sequence past the kinase domain at the C-end are more divergent between the two genes.

Figure 4.16 Reconstructed full-length cDNA of the *Bvcrk1* gene.

The boxed regions of the sequence represent the newly identified cDNA; the translation start site and the stop site are highlighted in gray colour. The black lines on the side of the sequence indicate the size of the two cRACE products. Thick black line- second 5′- cRACE product (1 to 886 bp); thin black line - first 5′-cRACE product (positions 500 To 1407). The 3′cDNA extension begins at position 2091.

4.5.5 Analysis of the predicted CRK1 peptide

The predicted protein sequence of the *crk1* gene was used as a query in BLAST and TFASTA homology searches. The *crk1* gene encodes a protein with structural similarities to serine/threonine kinases, and more specifically to number of CDK-like family members. The strong homology of the *Bvcrk1* with other CDKs is over the core catalytic domain (subdomains I - XI) of the protein kinases (figure 4.18).

It possesses an ATP-binding site GXGXXS (S substitutes the last G in the Rossmann motif) and Lys33, Asp127, Lys129, Asn132 and Asp145 (numbering according to Hscdk2) thought absolutely necessary in the orientation of the ATP and the transfer of the γ-phosphate of the ATP to the acceptor residue of the peptide substrate (Knighton *et al*, 1991; De Bondt *et al*, 1993).

The inhibitory phosphorylation sites at the N-terminus of the kinase catalytic domain present in the CDKs, Thr 14, and Tyr 15 (numbering according the Hscdk2, De Bondt *et al* 1993) exist also on the CRK1. Thr 160, necessary for the full activation of the kinase activity of the CDK complexes is conserved as well.

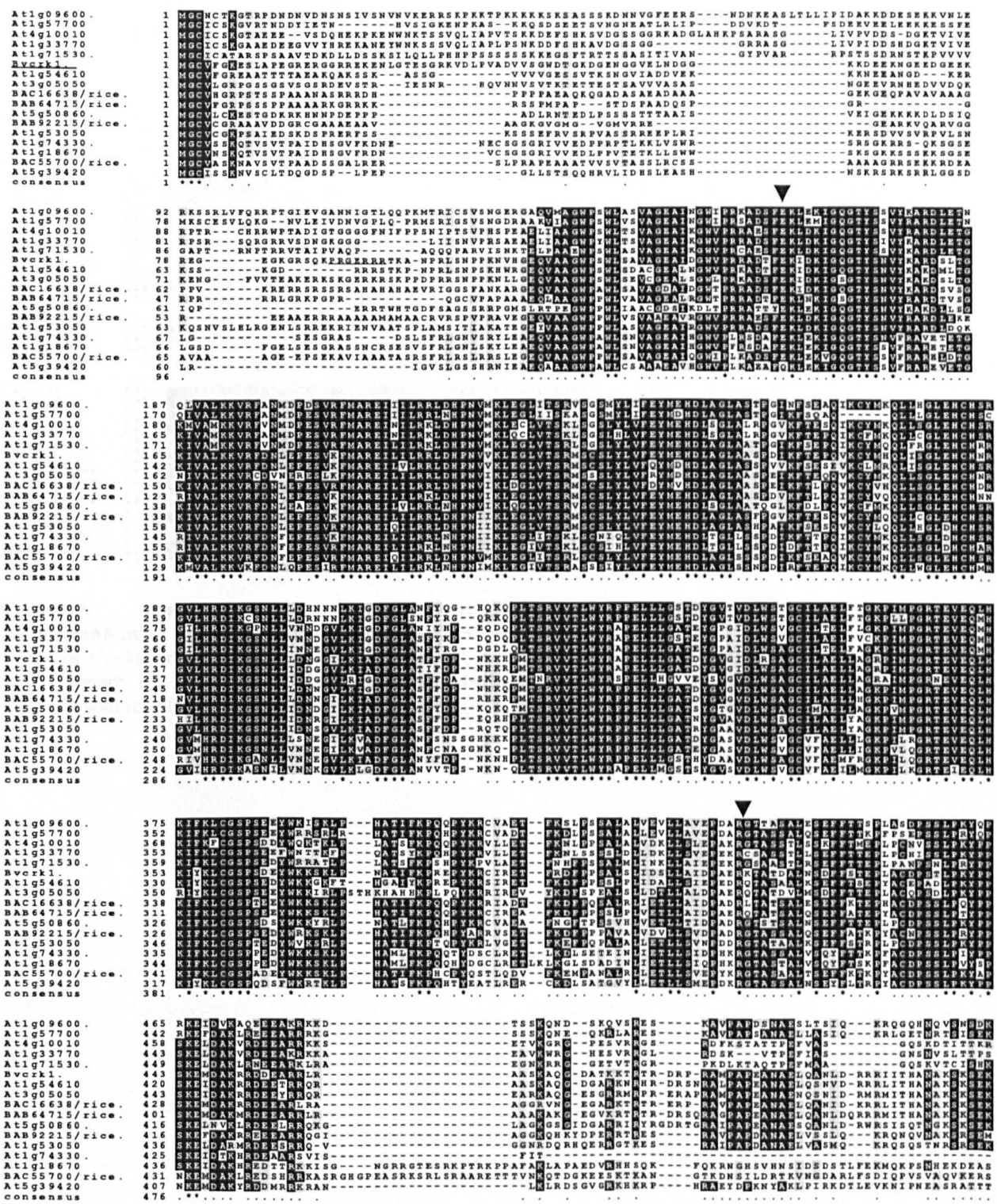
On the basis of sequence homology, CRK1 is closer to non-PSTAIRE CDKs like for instance the human HsPITAIRE, HsCdk7 (NRTALRE; accession No p50613), and HsCHED (PISSLRE; accession No s49330) or the rice OsCAK - NFTALRE (accession No p29620) and the Medicago Medsa;CDKC;1 (CAA65979). The sequence identity over the catalytic domain varies between 37.5% to 45.5% compared with PSTAIRE proteins where the identity is generally lower: 32%-35%.

One of the most divergent regions within the CRK1 catalytic domain is in the so-called cyclin-binding domain where the PSTAIRE motif appears as a "KFMARE". It differs from the majority of CDKs because it lacks an amino acid: Ile (in the PSTAIRE-type CDKs) or Leu present in the majority of non-PSTAIRE epitopes (e.g. HsPITALRE). The PSTAIRE region has been defined as important for the cyclin binding in the formation of an active kinase complex (figure 4.18).

The CRK1 peptide contains two possible strong NLSs (nuclear localization signals) (Boulikas, 1996). One is at the N-terminus (**PRGERRR**), the other at the C-terminus (**PRRKKK**) (figure 4.16). The signals were predicted by the PSORT programme available through the Bioinformatics services at MRC, HGMP, Hinxton, UK.Within the Bvcrk1-like group, most entries seem to have one, or two NLS which makes the group predpminantly nuclear proteins.

The homology data searches revealed that CRK1 is a member of a group of CDK-like proteins with a common FMARE motif and overall sequence homology higher than any other CDKs. For instance one of the closest homologues, At1g54610 has 63.5% sequence identity over the full length of the protein and 83.5% over the kinase catalytic domain. A multiple sequence alignment of the *crk1* homologues is shown in figure 4.16. It can be noted that the homology within this group of proteins exceeds the kinase catalytic domain. Nevertheless, the homology outside of the catalytic domain is lost towards the N- and C-termini of the proteins. It can be suggested that some functional specificity may be defined from these divergent parts of the CRK1-like proteins. It appears that up to the present date, the *Bvcrk1* gene is the most closely investigated member of this group of protein kinases. The relationship between the CRK1-like proteins and the remaining plant CDKs is given with the phylogenetic tree on figure 4.17. From this tree graph it appears that the CRK1-like sequences are more closely related to the plant CDK C-type kinases then to any other group. Within the CRK1 group, three distinct subgroups are formed due to the divergence ot the N-and C-terminal extensions (subgroups presented with different colours in figure 4.17). Overall, it can be concluded that the newly described subfamily of CRK1-like proeteins remains quite enigmatic as questions about the function of these proteins cannot be answered before performing further experiments to demonstrate their kinase activity and if they need cyclin partners for activity.

Figure 4.16 (for explanation see the end of the alignment on next page)



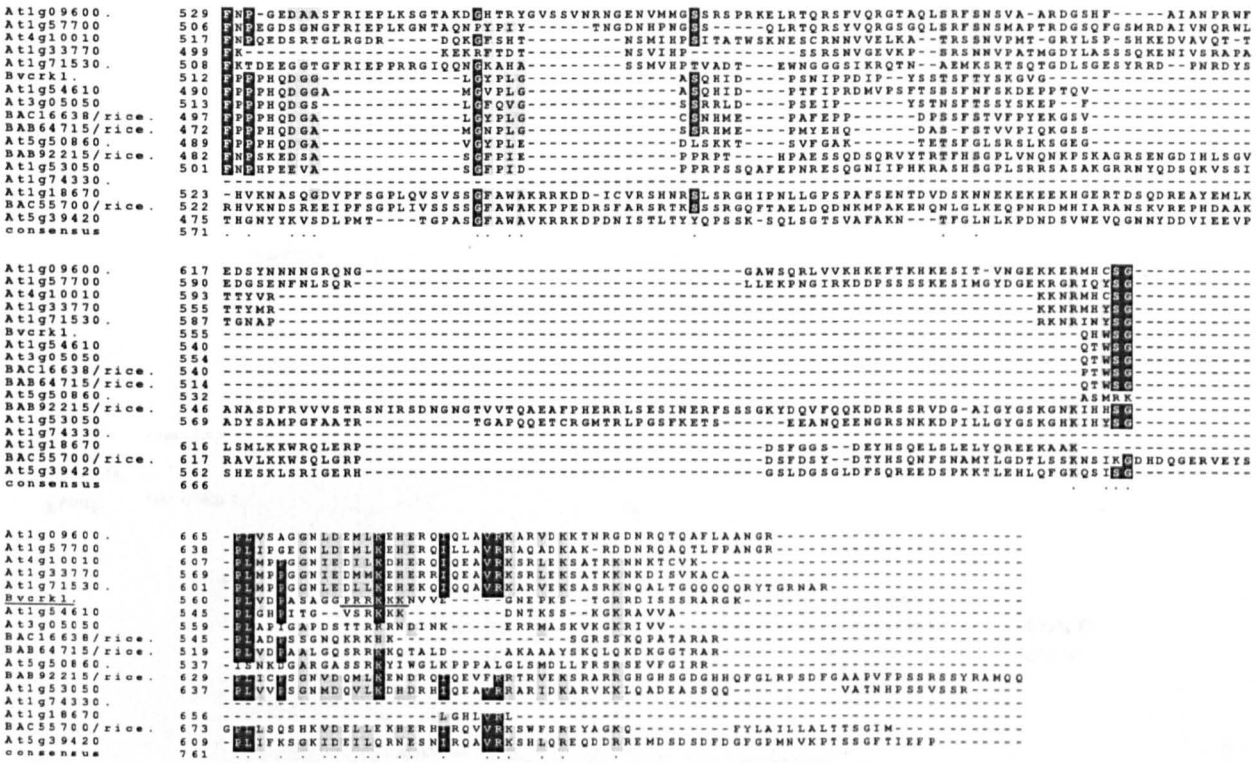


Figure 4.16 **A multiple sequence alignment of the *Bvcrk1*-like group of protein kinases**. The shaded box was drawn using the boxshade drawing facility of the http://www.blocks.fhcrc.org/blocks/process_blocks.html. A consensus is calculated for 0.55 of the sequences in the alignment. The *Bvcrk1* name is underlined for easier location, and the two NLS of the *Bvcrk1* are also underlined with red. The arrowheads indicate the start and the end of the kinase catalytic domain.

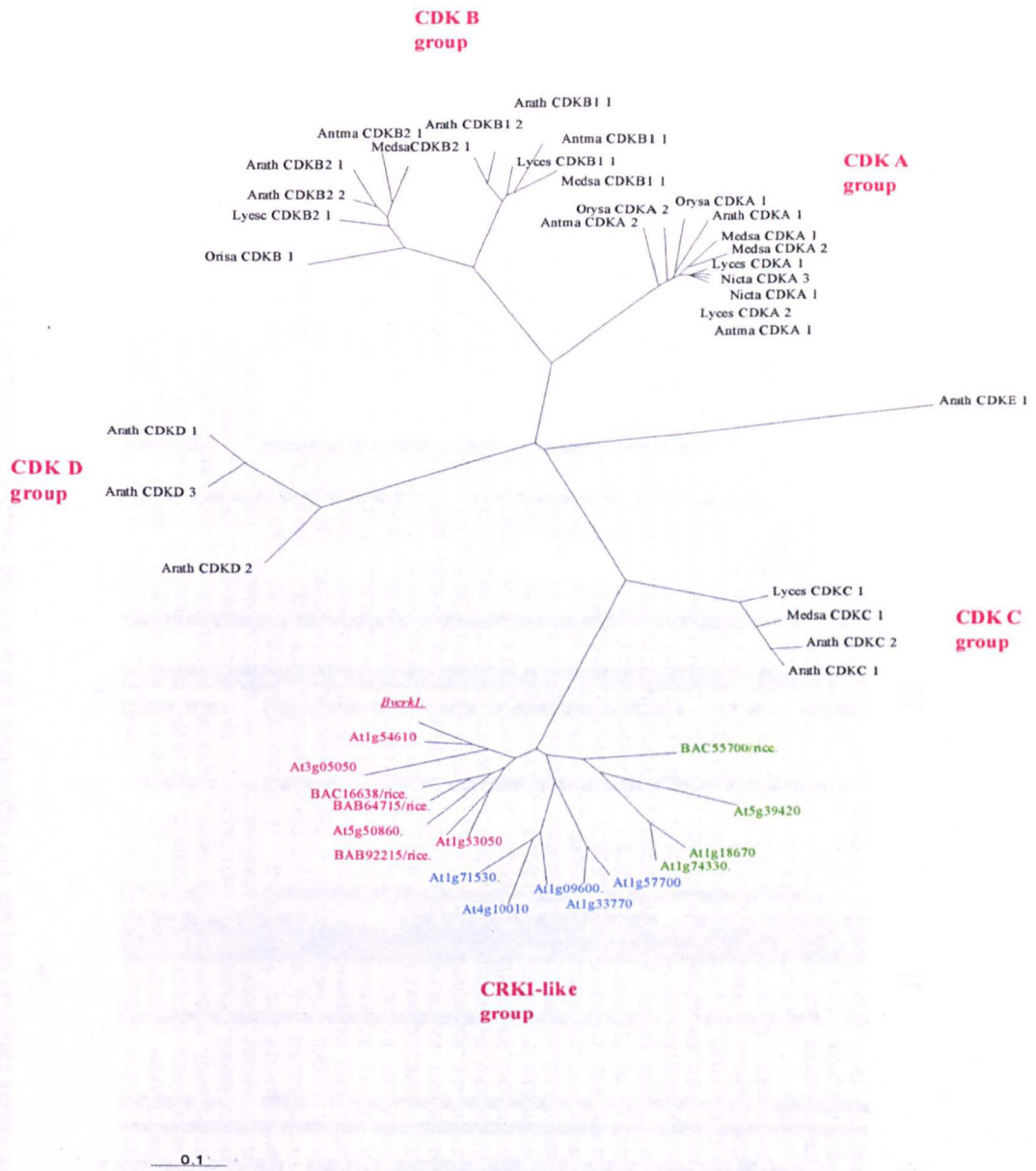
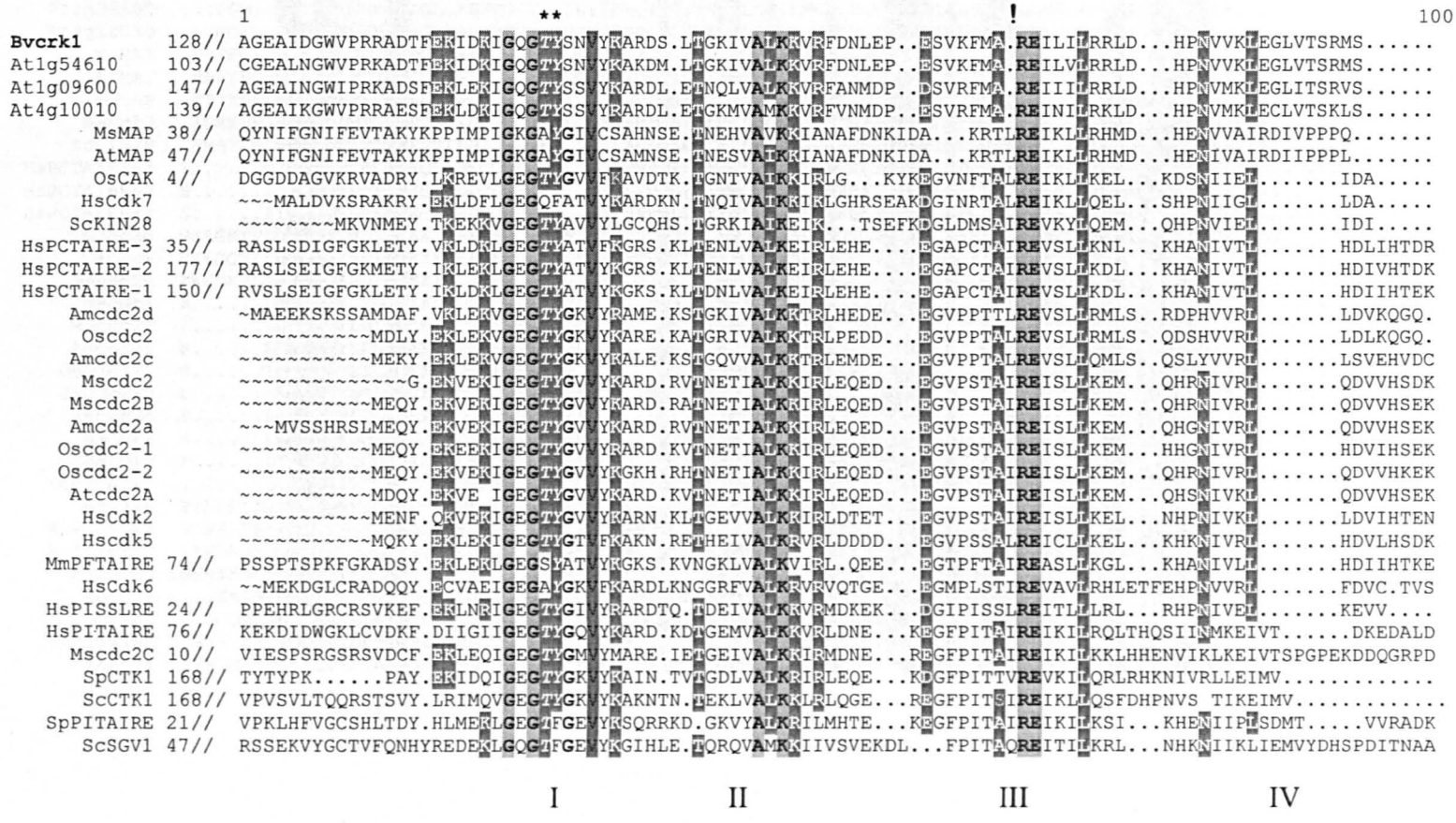


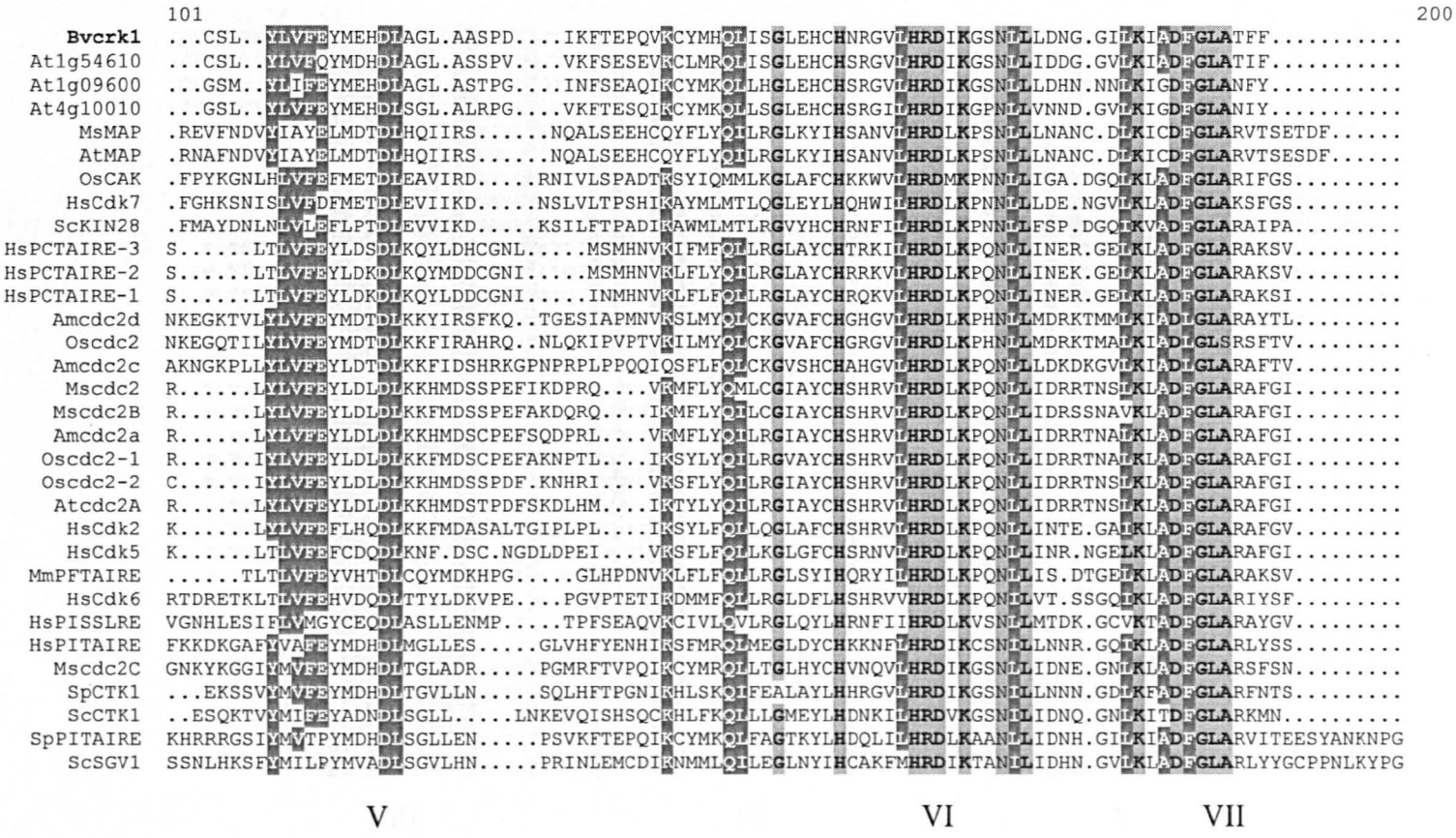
Figure 4.18 A radial tree of the plant CDKs, including the *crk1*-like group presented here in three different colours in accordance with the three subgroups existing within this group.

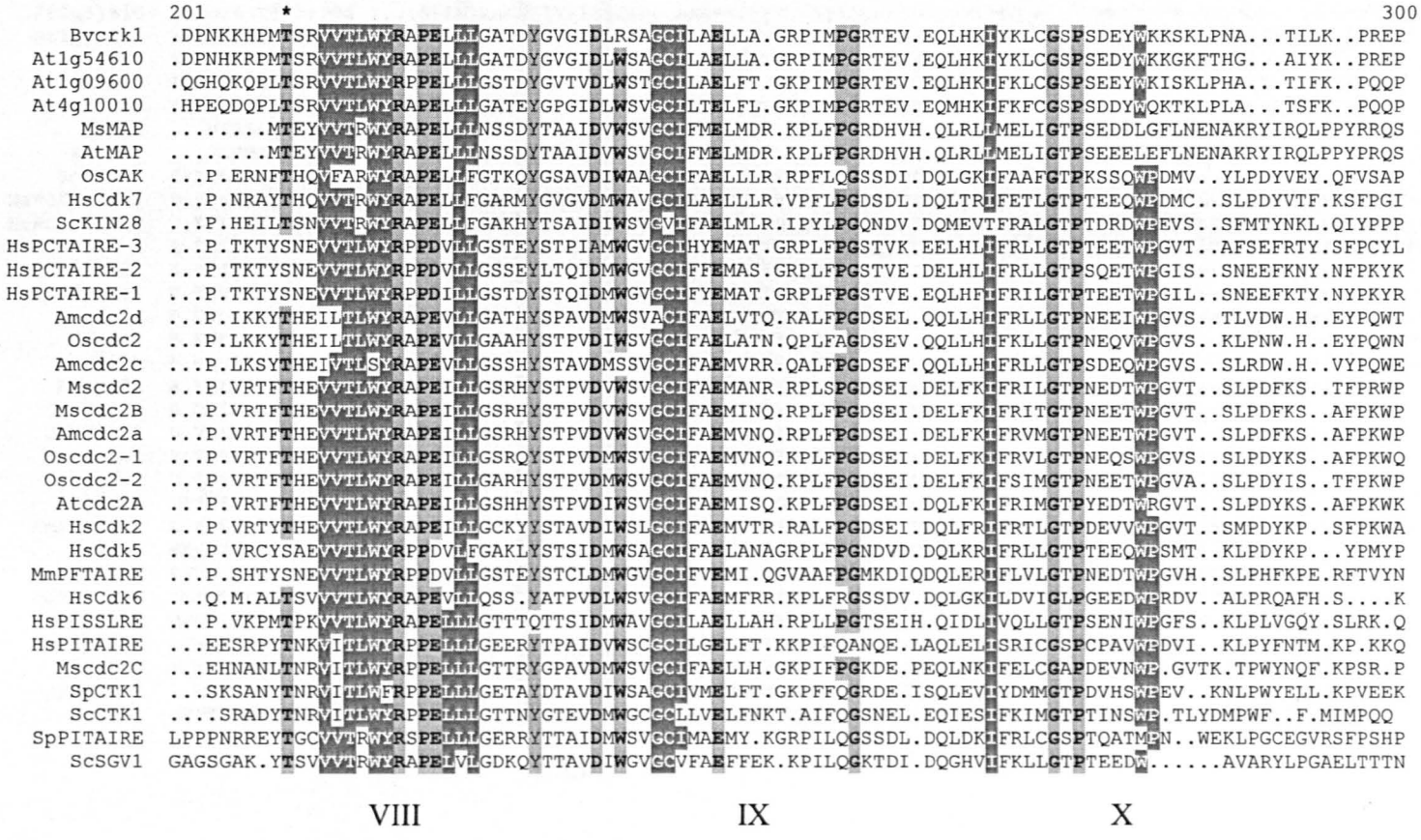
The tree was produced using a CLUSTAW generated Multiple Sequence Format alignment (http://www.ebi.ac.uk/clustalw/index.html) which was submitted to the BLOCKS MULTIPLE ALIGNMENT PROCESSOR (http://www.blocks.fhcrc.org/blocksl) to generate a Newick formatted tree file (http://www.proweb.org/proweb-bin/trees.cgi) and was then viewed with the Tree View software, version 1.6.6. The gene or accession numbers are given for the entries that shown by name only:

Bvcrk1/CAB89665; Arath CDKA;1/At3g48750; Antma;CDKA;1/CAA66233; Nicta;CDKA;1/AAB02567; Nicta;CDKA;2/AAB02568; Nicta;CDKA;3/BAA09369; Medsa;CDKA;2/CAA50038; Medsa;CDKA;1/AAB41817; Lyces;CDKA;1/CAA76700; Lyces;CDKA;2/CAA76701; Orysa;CDKA;2/CAA42923; Orysa;CDKA;1/CAA42922; Lyces;CDKB1;1/CAC15503; Lyesc;CDKB2;1/CAC15504; Orisa;CDKB;1/BAA19553; Medsa;CDKB1;1/CAA65980; MedsaCDKB2;1/CAA65982; Antma;CDKA;2/CAA66234; Antma;CDKB1;1/CAA66235; Antma;CDKB2;1/CAA66236; Arath;CDKC;2/At5g64960.; Arath;CDKB1;2/At2g38620; Arath;CDKB2;1/At1g76540; Arath;CDKB2;2/At1g20930; Arath;CDKD;1/At1g73690; Arath;CDKD;2/At1g66750; Arath;CDKD;3/At1g18040; Lyces;CDKC;1/CAC51391; Arath;CDKC;1/At5g10270; Medsa;CDKC;1/CAA65979; and Arath;CDKE;1/At5g63610.

Figure 4.18 **Multiple sequence alignment of the kinase catalytic domain.** For explanation and accession numbers see footnote 1 at the end of the alignment




101
200
Bvcrk1 ...CSL..YLVFEYMEHDLAGL.AASPD....IKFTEPQVKCYMHQLISGLEHCHNRGVLHRDIKGSNLLLDNG.GILKIADFGLATFF.....
At1g54610 ...CSL..YLVEQYMDHDLAGL.ASSPV....VKFSESEVKCLMRQLISGLEHCHSRGVLHRDIKGSNLLIDDG.GVLKIADFGLATIF.....
At1g09600 ...GSM..YLIFEYMEHDLAGL.ASTPG....INFSEAQIKCYMKQLLHGLEHCHSRGVLHRDIKGSNLLLDHN.NNLKIGDFGLANFY.....
At4g10010 ...GSL..YLVFEYMEHDLSGL.ALRPG....VKFTESQIKCYMKQLLSGLEHCHSRGILHRDIKGPNLLVNND.GVLKIGDFGLANIY.....
MsMAP .REVFNDVYIAYELMDTDLHQIIRS......NQALSEEHCQYFLYQILRGLKYIHSANVLHRDLKPSNLLLNANC.DLKICDFGLARVTSETDF.....
AtMAP .RNAFNDVYIAYELMDTDLHQIIRS......NQALSEEHCQYFLYQILRGLKYIHSANVLHRDLKPSNLLLNANC.DLKICDFGLARVTSESDF.....
OsCAK .FPYKGNLHLVFEFMETDLEAVIRD.....RNIVLSPADTKSYIQMMLKGLAFCHKKWVLHRDMKPNNLLIGA.DGQLKLADFGLARIFGS.....
HsCdk7 .FGHKSNISLVFDFMETDLEVIIKD.....NSLVLTPSHIKAYMLMTLQGLEYLHQHWILHRDLKPNNLLLDE.NGVLKLADFGLAKSFGS.....
ScKIN28 .FMAYDNLNLVLEFLPTDLEVVIKD.....KSILFTPADIKAWMLMTLRGVYHCHRNFILHRDLKPNNLLFSP.DGQIKVADFGLARAIPA.....
HsPCTAIRE-3 S......LTLVFEYLDSDLKQYLDHCGNL.....MSMHNVKIFMFQLLRGLAYCHTRKILHRDLKPQNLLINER.GELKLADFGLARAKSV.....
HsPCTAIRE-2 S......LTLVFEYLDKDLKQYMDDCGNI.....MSMHNVKLFLYQILRGLAYCHRRKVLHRDLKPQNLLINEK.GELKLADFGLARAKSV.....
HsPCTAIRE-1 S......LTLVFEYLDKDLKQYLDDCGNI.....INMHNVKLFLFQLLRGLAYCHRQKVLHRDLKPQNLLINER.GELKLADFGLARAKSI.....
Amcdc2d NKEGKTVLYLVFEYMDTDLKKYIRSFKQ..TGESIAPMNVKSLMYQLCKGVAFCHGHGVLHRDLKPHNLLMDRKTMMLKIADLGLARAYTL.....
Oscdc2 NKEGQTILYLVFEYMDTDLKKFIRAHRQ..NLQKIPVPTVKILMYQLCKGVAFCHGRGVLHRDLKPHNLLMDRKTMALKIADLGLSRSFTV.....
Amcdc2c AKNGKPLLYLVFEYLDTDLKKFIDSHRKGPNPRPLPPQQIQSFLFQLCKGVSHCHAHGVLHRDLKPQNLLLDKDKGVLKIADLGLARAFTV.....
Mscdc2 R......LYLVFEYLDLDLKKHMDSSPEFIKDPRQ....VKMFLYQMLCGIAYCHSHRVLHRDLKPQNLLIDRRTNSLKLADFGLARAFGI.....
Mscdc2B R......LYLVFEYLDLDLKKFMDSSPEFAKDQRQ....IKMFLYQILCGIAYCHSHRVLHRDLKPQNLLIDRSSNAVKLADFGLARAFGI.....
Amcdc2a R......LYLVFEYLDLDLKKHMDSCPEFSQDPRL....VKMFLYQILRGIAYCHSHRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGI.....
Oscdc2-1 R......IYLVFEYLDLDLKKFMDSCPEFAKNPTL....IKSYLYQILRGVAYCHSHRVLHRDLKPQNLLIDRRTNALKLADFGLARAFGI.....
Oscdc2-2 C......IYLVFEYLDLDLKKHMDSSPDF.KNHRI....VKSFLYQILRGIAYCHSHRVLHRDLKPQNLLIDRRTNSLKLADFGLARAFGI.....
Atcdc2A R......LYLVFEYLDLDLKKHMDSTPDFSKDLHM....IKTYLYQILRGIAYCHSHRVLHRDLKPQNLLIDRRTNSLKLADFGLARAFGI.....
HsCdk2 K......LYLVFEFLHQDLKKFMDASALTGIPLPL....IKSYLFQLLQGLAFCHSHRVLHRDLKPQNLLINTE.GAIKLADFGLARAFGV.....
HsCdk5 K......LTLVFEFCDQDLKNF.DSC.NGDLDPEI....VKSFLFQLLKGLGFCHSRNVLHRDLKPQNLLINR.NGELKLADFGLARAFGI.....
MmPFTAIRE ......TLTLVFEYVHTDLCQYMDKHPG.....GLHPDNVKLFLFQLLRGLSYIHQRYILHRDLKPQNLLIS.DTGELKLADFGLARAKSV.....
HsCdk6 RTDRETKLTLVFEHVDQDLTTYLDKVPE....PGVPTETIKDMMFQLLRGLDFLHSHRVVHRDLKPQNILVT.SSGQIKLADFGLARIYSF.....
HsPISSLRE VGNHLESIFLVMGYCEQDLASLLENMP.....TPFSEAQVKCIVLQVLRGLQYLHRNFIIHRDLKVSNLLMTDK.GCVKTADFGLARAYGV.....
HsPITAIRE FKKDKGAFYVAFEYMDHDLMGLLES.....GLVHFYENHIKSFMRQLMEGLDYCHKKNFLHRDIKCSNLLLNNR.GQIKLADFGLARLYSS.....
Mscdc2C GNKYKGGIYMVFEYMDHDLTGLADR.....PGMRFTVPQIKCYMRQLLTGLHYCHVNQVLHRDIKGSNLLIDNE.GNLKLADFGLARSFSN.....
SpCTK1 ...EKSSVYMVFEYMDHDLTGVLLN.....SQLHFTPGNIKHLSKQIFEALAYLHHRGVLHRDIKGSNILLNNN.GDLKFADFGLARFNTS.....
ScCTK1 ..ESQKTVYMIFEYADNDLSGLL.....LNKEVQISHSQCKHLFKQLLLGMEYLHDNKILHRDVKGSNILIDNQ.GNLKITDFGLARKMN.....
SpPITAIRE KHRRRGSIYMVTPYMDHDLSGLLEN.....PSVKFTEPQIKCYMKQLFAGTKYLHDQLILHRDLKAANLLIDNH.GILKIADFGLARVITEESYANKNPG
ScSGV1 SSNLHKSFYMILPYMVADLSGVLHN.....PRINLEMCDIKNMMLQILEGLNYIHCAKFMHRDIKTANILIDHN.GVLKLADFGLARLYYGCPPNLKYPG
V
VI
VII


201
300
Bvcrk1 .DPNKKHPMTSRVVTLWYRAPELLLGATDYGVGIDLRSAGCILAELLA.GRPIMPGRTEV.EQLHKIYKLCGSPSDEYWKKSKLPNA...TILK..PREP
At1g54610 .DPNHKRPMTSRVVTLWYRAPELLLGATDYGVGIDLWSAGCILAELLA.GRPIMPGRTEV.EQLHKIYKLCGSPSEDYWKKGKFTHG...AIYK..PREP
At1g09600 .QGHQKQPLTSRVVTLWYRPPELLLGSTDYGVTVDLWSTGCILAELFT.GKPIMPGRTEV.EQLHKIFKLCGSPSEEYWKISKLPHA...TIFK..PQQP
At4g10010 .HPEQDQPLTSRVVTLWYRAPELLLGATEYGPGIDLWSVGCILTELFL.GKPIMPGRTEV.EQMHKIFKFCGSPSDDYWQKTKLPLA...TSFK..PQQP
MsMAP ........MTEYVVTRWYRAPELLLNSSDYTAAIDVWSVGCIFMELMDR.KPLFPGRDHVH.QLRLLMELIGTPSEDDLGFLNENAKRYIRQLPPYRRQS
AtMAP ........MTEYVVTRWYRAPELLLNSSDYTAAIDVWSVGCIFMELMDR.KPLFPGRDHVH.QLRLLMELIGTPSEEELEFLNENAKRYIRQLPPYPRQS
OsCAK ...P.ERNFTHQVFARWYRAPELLFGTKQYGSAVDIWAAGCIFAELLLR.RPFLQGSSDI.DQLGKIFAAFGTPKSSQWPDMV..YLPDYVEY.QFVSAP
HsCdk7 ...P.NRAYTHQVVTRWYRAPELLFGARMYGVGVDMWAVGCILAELLLR.VPFLPGDSDL.DQLTRIFETLGTPTEEQWPDMC..SLPDYVTF.KSFPGI
ScKIN28 ...P.HEILTSNVVTRWYRAPELLFGAKHYTSAIDIWSVGVIFAELMLR.IPYLPGQNDV.DQMEVTFRALGTPTDRDWPEVS..SFMTYNKL.QIYPPP
HsPCTAIRE-3 ...P.TKTYSNEVVTLWYRPPDVLLGSTEYSTPIAMWGVGCIHYEMAT.GRPLFPGSTVK.EELHLIFRLLGTPTEETWPGVT..AFSEFRTY.SFPCYL
HsPCTAIRE-2 ...P.TKTYSNEVVTLWYRPPDVLLGSSEYLTQIDMWGVGCIFFEMAS.GRPLFPGSTVE.DELHLIFRLLGTPSQETWPGIS..SNEEFKNY.NFPKYK
HsPCTAIRE-1 ...P.TKTYSNEVVTLWYRPPDILLGSTDYSTQIDMWGVGCIFYEMAT.GRPLFPGSTVE.EQLHFIFRILGTPTEETWPGIL..SNEEFKTY.NYPKYR
Amcdc2d ...P.IKKYTHEILTLWYRAPEVLLGATHYSPAVDMWSVACIFAELVTQ.KALFPGDSEL.QQLLHIFRLLGTPNEEIWPGVS..TLVDW.H..EYPQWT
Oscdc2 ...P.LKKYTHEILTLWYRAPEVLLGAAHYSTPVDIWSVGCIFAELATN.QPLFAGDSEV.QQLLHIFKLLGTPNEQVWPGVS..KLPNW.H..EYPQWN
Amcdc2c ...P.LKSYTHEIVTLSYRAPEVLLGSSHYSTAVDMSSVGCIFAEMVRR.QALFPGDSEF.QQLLHIFRLLGTPSDEQWPGVS..SLRDW.H..VYPQWE
Mscdc2 ...P.VRTFTHEVVTLWYRAPEILLGSRHYSTPVDVWSVGCIFAEMANR.RPLSPGDSEI.DELFKIFRILGTPNEDTWPGVT..SLPDFKS..TFPRWP
Mscdc2B ...P.VRTFTHEVVTLWYRAPEILLGSRHYSTPVDVWSVGCIFAEMINQ.RPLFPGDSEI.DELFKIFRITGTPNEETWPGVT..SLPDFKS..AFPKWP
Amcdc2a ...P.VRTFTHEVVTLWYRAPEILLGSRHYSTPVDVWSVGCIFAEMVNQ.RPLFPGDSEI.DELFKIFRVMGTPNEETWPGVT..SLPDFKS..AFPKWP
Oscdc2-1 ...P.VRTFTHEVVTLWYRAPEILLGSRQYSTPVDMWSVGCIFAEMVNQ.KPLFPGDSEI.DELFKIFRVLGTPNEQSWPGVS..SLPDYKS..AFPKWQ
Oscdc2-2 ...P.VRTFTHEVVTLWYRAPEILLGARHYSTPVDMWSVGCIFAEMVNQ.KPLFPGDSEI.DELFKIFSIMGTPNEETWPGVA..SLPDYIS..TFPKWP
Atcdc2A ...P.VRTFTHEVVTLWYRAPEILLGSHHYSTPVDIWSVGCIFAEMISQ.KPLFPGDSEI.DQLFKIFRIMGTPYEDTWRGVT..SLPDYKS..AFPKWK
HsCdk2 ...P.VRTYTHEVVTLWYRAPEILLGCKYYSTAVDIWSLGCIFAEMVTR.RALFPGDSEI.DQLFRIFRTLGTPDEVVWPGVT..SMPDYKP..SFPKWA
HsCdk5 ...P.VRCYSAEVVTLWYRPPDVFFGAKLYSTSIDMWSAGCIFAELANAGRPLFPGNDVD.DQLKRIFRLLGTPTEEQWPSMT..KLPDYKP...YPMYP
MmPFTAIRE ...P.SHTYSNEVVTLWYRPPDVLLGSTEYSTCLDMWGVGCIFVEMI.QGVAAFPGMKDIQDQLERIFLVLGTPNEDTWPGVH..SLPHFKPE.RFTVYN
HsCdk6 ...Q.M.ALTSVVVTLWYRAPEVLLQSS.YATPVDLWSVGCIFAEMFRR.KPLFRGSSDV.DQLGKILDVIGLPGEEDWPRDV..ALPRQAFH.S....K
HsPISSLRE ...P.VKPMTPKVVTLWYRAPELLLGTTTQTTSIDMWAVGCILAELLAH.RPLLPGTSEIH.QIDLIVQLLGTPSENIWPGFS..KLPLVGQY.SLRK.Q
HsPITAIRE ...EESRPYTNKVITLWYRPPELLLGEERYTPAIDVWSCGCILGELFT.KKPIFQANQE.LAQLELISRICGSPCPAVWPDVI..KLPYFNTM.KP.KKQ
Mscdc2C ...EHNANLTNRVITLWYRPPELLLGTTRYGPAVDMWSVGCIFAELLH.GKPIFPGKDE.PEQLNKIFELCGAPDEVNWP.GVTK.TPWYNQF.KPSR.P
SpCTK1 ...SKSANYTNRVITLWFRPPELLLGETAYDTAVDIWSAGCIVMELFT.GKPFFQGRDE.ISQLEVIYDMMGTPDVHSWPEV..KNLPWYELL.KPVEEK
ScCTK1 ....SRADYTNRVITLWYRPPELLLGTTNYGTEVDMWGCGCLLVELFNKT.AIFQGSNEL.EQIESIFKIMGTPTINSWP.TLYDMPWF..F.MIMPQQ
SpPITAIRE LPPPNRREYTGCVVTRWYRSPELLLGERRYTTAIDMWSVGCIMAEMY.KGRPILQGSSDL.DQLDKIFRLCGSPTQATMPN..WEKLPGCEGVRSFPSHP
ScSGV1 GAGSGAK.YTSVVVTRWYRAPELVLGDKQYTTAVDIWGVGCVFAEFFEK.KPILQGKTDI.DQGHVIFKLLGTPTEEDW......AVARYLPGAELTTTN
VIII
IX
X

[1] Identical amino-acid residues are shown as black bold letters on gray background. The nearly conserved amino acid residues are given as white letters on darker gray background. Roman numerals indicate the 11 conserved kinase domains (Hanks *et al.*, 1992). The phosphorylation sites are marked with an asterisk (*) and represent Thr 14, Tyr 15 and Thr 160 according to the numbering of Hscdk2. The missing Ile or Leu in the "PSTAIRE" region in some of the kinases is marked with an exclamation mark (!) above the alignment.

Accession numbers of the sequences in the multiple alignment: Bvcrk1 (CCCAAAB89665; (p24923); Mscdc2B (q05006); Mscdc2C (p93320); Mscdc2D; MsMAP kinase (q07176); AtMAP kinase (q39026); Oscdc2 (q40734); Oscdcd-1 (p29618); Oscdc2-2 (p29619); OsCAK (p29620); Amcdc2a (x97637); Amcdc2c (x97639); Amcdc2d (x97640); Hscdk2 (p24941); Hscdk5 (q00535); Hscdk6 (q00534); Hscdk7 (p50613); HsPCTAIRE-1 (Q00536); HsPCTAIRE-2 (Q00537); HsPCTAIRE-3 (Q07002); HsPISSLRE (s49330); HsPITAIRE ( ); MmPFTAIRE (u62391); ScKIN28 (A25698); ScCTK1 (q03957); ScSGV1 (a39526); SpCTK1 (z99165); SpPITAIRE (ab00453413); At-Arabidopsis thaliana; Ms -Medicago sativa; Os - Orisa sativa; Am - Antherhynuim majum; Hs - Homo sapiens; Mm - Mus musculus; Sc - Saccharomyces cerevise; Sp - Schizosacchoaromyces pombe.

PAGE NUMBERING AS ORIGINAL

CHAPTER 5 EXPRESSION OF *Bvcrk1* GENE

The information presented in this chapter was obtained using the *in situ* hybridisation method to detect the expression pattern of the *Bvcrk1* gene in different tissues of sugar beet.

Originally, the idea was to use different time points in the development of the plants as well as different tissues from each time point in order to determine the activity of the *Bvcrk1* gene. These included the seedling, young plants of 1 and 3 true leaves (the latter about one month old), and older plants (3 months old). The ages were selected to match the presumed stages in the tap root formation and development because of the emphasis or the special interest in this organ, and in particular in the formation and development of the multiple cambium layers. However, this plan was not followed completely due to some difficulties in the sample preparation.

The interest in *crk1* expression pattern came from some earlier results, which suggested for a role of the gene in the early stages of the cell cycle. In particular it was shown that the expression of the gene was induced by sucrose, kinetin and IAA, as well as by MS macronutrients in quiescent cell suspension cultures of sugar beet cells. Upon transfer to fresh culture medium transcripts of *Bvcrk1* were detectable before the expression of histone H4 and Artha;*Cyc B1;1*. (Fowler *et al.*, 1998).

The *in situ* hybridisations were seen as the next step in the characterisation of the expression of the gene. It was interesting to see where in the tissue transcripts could be detected and whether the gene showed any specificity in time and space.

5.1 Materials and methods

5.1.1 Plant material

Sugar beet plants, cultivar Roberta were used for the *in situ* hybridisations. The plants were grown in individual pots with sterilised compost in a growth room with 16 hours light and 24°C temperature regime maintained throughout. The plants were regularly transferred to bigger sized pots so that the growth of the root was not restricted by the size of the pot.

For the collection of seedlings, seeds were first soaked from 2 to 16 hours in water (this procedure speeded up the germination of the seeds with 4 to 5 days). Seeds were germinated in trays on moist tissue paper in growth chamber with 16 hours light and temperature 24°C. The germination process normally took about 2 to 4 days.

A list of the plant tissues used and their ages is given below for clarification.

1. Root tip sections, taken from 2 to 4 day old seedlings. The first 2 mm from the root tip were used.
2. Root sections, taken from young seedlings as in 1, but at 5 to 8 mm from the root tip.
3. Apical meristem sections, taken from seedlings and plants with one true leaf emerging (about 10 days old).
4. Apical meristem sections of 3 months old plants. The leaf bud area was taken as a whole and immersed immediately in fixative.
5. Root sections, taken from 3-month-old sugar beet plants. 2 to 3 mm thick sections were made with a hand razor and fixed.

5.1.2 Preparation of the plant material

- Fixation

The plant tissue was cut into small sections about 1 to 2 mm thick with a razor blade or with a hand microtome in a petri dish with some 1 x PBS. The sections were transferred to glass vials with fixative (4% paraformaldehyde in 0.01 M phosphate buffer). A vacuum was applied to the sections for the first hour of fixation in order to improve the penetration of fixative into the tissue. The fixative was replaced with fresh solution and the sections were left at 4°C overnight.

- Dehydration

The fixative was removed and the sections were passed through the following steps for dehydration: 1x buffer (10 mM sodium phosphate, pH 6.8, 100 mM NaCl), sterile distilled water, 10%, 30%, 50%, 70%, 80%, 90% for 20 minutes each at room temperature, and finally 100% ethanol three times for 1 hour also at room temperature. The sections were left in 100% ethanol overnight at 4°C.

- Xylene infiltration

The 100% ethanol was replaced by mixtures of xylene and ethanol in ratio of 1:3, 1:1, 3:1 for one hour each at room temperature. Finally two changes of 100% xylene were made and the sections were left in the xylene at 4°C overnight.

- Parafin infiltration

Paraplast chips (Paraplast Embedding Media, Sigma) were added one every hour until they reached half the volume of the xylene (15 to 20 chips). At this stage the sections were left at room temperature overnight and subsequently transferred to 42°C overnight to dissolve the paraplast completely. The vials were moved to 60°C oven and the xylene/paraplast mixture was exchanged with molten filtered paraplast. The paraplast was changed twice or three times a day for the next three days.

- Embedding

The sections were embedded in parapalst by carefully pouring the contents of the vials into trays containing small amount of melted paraplast at 60°C. With the help of a heated needle the sections were positioned in the desired orientation and then the mould was transferred into ice cold water until the blocks set completely. The paraplast blocks were stored at 4°C.

- Sectioning and removal of the parafin

7 to 10 μm parafin sections were cut on a rotating microtome (HM 335E, Microtim) with a steel knife. The ribbons were placed on clean paper and cut to fragments about 5 cm long. Selected ribbons were stretched with sterile water on poly-L-lysine coated glass slides at 42°C for up to a minute. The water was drained off as thoroughly as possible and the slides were left to dry for 4 hours to overnight on a 42°C hot plate (SH3D, Stuart Scientific, UK). The slides were transferred to 25 slot glass racks and immersed in 100 % xylene for 15 minutes at room temperature to remove the parapalst. The slides were moved to fresh 100% xylene for another 15 minutes and then passed through 50% xylene in 100% ethanol for 20 minutes. Finally they were passed in 100% ethanol for 10 minutes to remove the remaining xylene. The slides were air-dried and examined under a microscope to select suitable sections for *in situ* hybridisation.

The chosen slides were put in a plastic box, sealed with tape and stored at -20°C until needed for the hybridisation.

5.1.3 Preparation of slides

SuperFrost glass slides (24mm x 50mm x 1mm, BDH Merck) were baked for 24 hours at 180°C to eliminate any possible RNase contamination and then coated with poly-L-lysine

(100µg/ml poly-L-lysine, (Sigma) solution in 10mM Tris-HCl (pH 8.0) for 15 minutes). Slides were allowed to air-dry and then were stored in boxes at -20°C. Prior to being opened again, the boxes with slides were left at room temperature for 2 to 3 hours to avoid condensation.

5.1.4 Probes for in situ hybridisation

- Probes used in the hybridisations

1. *Arabidopsis* Histone H4 230 bp cDNA clone in pBlueScript SK- was available in the laboratory and was a gift from Dr. C. Gigot in 1996 (Institute de Biologie Moleculaire des Palntes, Montpellier, France).
2. A 1.34 kb cDNA fragment of *Bvcrk1* cloned into pGEM-T vector (+ 671 bp to + 2031 bp on *Bvcrk1* cDNA).
3. Sugar beet *Betvu;CDKA;1*,a probe which was a 353 bp fragment (accession number Z71702) in pGEM-T vector.

- Probe preparation

The plasmids with the histone H4 and *crk1* and CDKA;1 inserts were linearised by cutting within the MCS of the respective plasmids (table 5.1) to produce the template for *in vitro* transcription in the sense and antisense orientations using the DIG RNA labelling Kit (Boeringer Mannheim).

Table 5.1 **Restriction enzymes used for the generation of the linear sense and antisense probes**

probe	antisense	sense
crk1 / p GEM-T	Nco I /SP6	Sal I / T7
At H4 /pBlueScript SK-	Not I / T7	Sal I / T3
Betvu;CDKA;1	Not I / T7	Sal I / T3

The digests were carried out for 2 hours. A third of the reaction was loaded on 0.8% 1xTAE agarose gel to check if the vector was linearised completely. The remaining two thirds of the reactions were purified by phenol, phenol/chloroform, and chloroform extraction and precipitated with ethanol. The DNA was re-dissolved in sterile distilled water at final concentration of 200 $ng.\mu l^{-1}$.

- Probe labelling by *in vitro* transcription

The linearised probes were labelled using DIG RNA labelling kit (Boehringer Mannheim) and following the instructions of the supplier's protocol:

Probe labelling	
components	**volume**
linearised template DNA	1μg (5μl)
10 x transcription buffer	2μl
NTP labelling mixture	2 μl
SP6, T7 or T3 RNA polymerase	2 μl
H_2O	9 μl
RNase inhibitor (20 U μl^{-1})	1 μl

The transcription was for 2 hours at 37°C, then the template DNA was removed by adding 2μl of RNase-free DNase I (10 U μl^{-1}), and incubating the reaction for 15 minutes at 37°C. The reaction was stopped with 2μl 0.2M EDTA, pH 8.0.

The content of the transcription reaction was passed through a ChromaSpin column-200 DEPC-H_2O, which purified the DIG-labelled riboprobes.

The transcripts were analysed by gel electrophoresis and ethidium bromide staining. One tenth of the purified RNA (7μl) was loaded on a 1.5% 1xTAE gel in order to check the transcription efficiency and the size of the RNA transcripts.

- Probe hydrolysis

The 1.34 kb probe of *crk 1* was hydrolysed as described by Cox and Goldberg in order to reduce the size of the fragments (to about 0.15 kb) so that they penetrate the sample's tissue and access the target mRNA sequences better. The time for hydrolysis was calculated as: $t = (Lo - Lf) \div (K \times Lo \times Lf)$, where t is the incubation time in minutes, Lo is the starting length in kb, Lf is the desired final length in kb, and K=0.11.

The hydrolysis was done by mixing 50 μl of the labelled RNA probe with 30μl 0.2M Na_2CO_3 and 20 μl 0.2M $NaHCO_3$ and incubating the mixture at 60°C for the calculated time (50 minutes for crk1). The hydrolysis reaction was stopped by adding 3μl 3M sodium acetate (pH 6) and 5μl 10% glacial acetic acid.

The RNA probe was purified through ChromaSpin column 100-DEPC H_2O and stored at minus 20°C before use.

5.2. IN SITU HYBRIDISATION

5.2. 1 Prehybridisation procedures

1. The slides selected for hybridisation were taken from -20°C and left in the storage box to adjust to room temperature before the hydration steps.

2. The slides were passed through the following ethanol series: 2 x 100% ethanol, then 90%, 70%, 50%, 10% ethanol, and three changes of sterile distilled water, 1 minute each step.

3. The slides were incubated in 100mM Tris-HCl (pH 7.5), 50mM EDTA (pH 8.0) buffer with 1μg/ml proteinase K (Sigma) at 37°C for 30 minutes.

4. The slides were washed with sterile distilled water 3 times, 1 minute each.

5. The slides were transferred to 0.1M triethanolamine, pH 8.0 (Sigma) at room temperature for 10 minutes. Acetic anhydride (Sigma) was added to final concentration of 0.25% and the slides were incubated for further 10 minutes.

6. Slides were dehydrated by using the same ethanol series (10, 30, 50, 70, 90, and 100%).

7. The sections were dried first at room temperature for 1 hour and then under vacuum for 1 to 2 hours followed immediately by the hybridisation step (5.2.2).

5.2.2 Hybridisation

- Hybridisation mix preparation

The hybridisation buffer contained the following components (in final concentrations):
50% double crystallised formamide, 300 mM NaCl, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% Ficoll, 0.02% polyvinylpyrolidone, 0.02% BSA, 10% Dextran sulfate, 60mM DTT, 500 μg ml^{-1} poly (A), 150 μg ml^{-1} yeast transfer RNA, and the RNA probe. The RNA probe concentration was kept at 100 ng per slide.

The components of the hybridisation buffer were divided into two solutions: A (kept at 37°C) and B, which were then mixed in ratio 9 to 1. Solution B which contained the probe was heated to 80°C for 5 minutes to denature the RNA and then added to solution A. One hundred to 120 μl of hybridisation buffer were applied to each slide. Coverslips were put over the hybridisation solution and the slides were left to incubate at 42°C overnight in humid slide boxes.

Hybridisation buffer	
Solution A (per slide)	**Solution B (per slide)**
- 300 µl formamide (double crystallised, see appendix C) - 6 µl 5 M NaCl - 10 µl 100 mM Tris-HCl, pH 7.5, 10 mM EDTA - 2 µl 50 x Denhardt's - 20 µl 50 % Dextran Sulfate (Wt 500,000 Sigma) - 1.5 µl 3 M DTT - 1.5µl H_2O	- 1 µl 3M DTT - 1.5 µl 10 µg μl^{-1} yeast tRNA (Sigma) - 2.5 µl 20 µg μl^{-1} polyA (Sigma) - 5 µl RNA probe and water

- Washes

After hybridisation, slides were placed in slide glass racks and passed through the following washing steps:

1. Cover slips were removed from slides by immersing the slide racks in a container with 4 x SSC solution for about 10 minutes until all of the cover slips fell from the slides.

2. Three changes of 4 x SSC solution, 10 minutes each at room temperature on a slow moving shaker.

3. The non-hybridised probe was removed by RNase A treatment (50 µg/ml RNase A (Sigma) in 500 mM NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 7.5). The slides were incubated with the RNase A for 30 minutes at 37°C followed by five 20 minute washes with the RNase buffer alone also at 37°C.

4. Finally, the slides were washed with 2 x SSC for 30 minutes at room temperature.

5.2.3 Signal detection

1. Slides were washed in PBT for 5 minutes.

2. Slides were incubated in 0.5% Blocking reagent (Boerhinger Mannheim) in PBT, for one hour at room temperature.

3. Slides were rinsed with PBT for 1 minute

4. Antibody incubation: 500 µl of antibody solution (500µl PBT pH 7.5, 10 µl BSA (10 mg ml^{-1}), 1:2000 of Anti-DIG-AP conjugate, Boehringer Mannheim) was applied to the sections. The slides were covered with coverslips and left for 1 hour in a dark humid box.

5. The slides were passed through three 5 minute washes in PBT, pH 7.5 at room temperature.

6. Slides were immersed in detection buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 5mM $MgCl_2$) for 5 minutes.

7. The slides were transferred in detection buffer containing 1 mM levamisole (Sigma) for 5 minutes in order to inhibit the endogenous phosphatases.

8. The solution was drained as thoroughly as possible from each slide. One hundred microlitres of detection buffer containing NBT/BCIP substrate was applied per each slide. 10 ml detection buffer was mixed with 45µl nitroblue tetrazoleum (NBT) stock (100 mg ml^{-1} in dimethylformamide from Boerhinger Mannheim) and 33µl 5-bromo-4-chloro-3-indolyl-phosphate, 4-toluidine salt (BCIP) stock (50 mg ml^{-1} in dimethylformamide, Boerhinger Mannheim). New coverslips were put over the sections to prevent the evaporation of the detection solution. The slides were left in the dark for about 12 hours before the detection reaction was stopped.

9. The slides were checked under microscope for the development of colour, and the reaction was stopped with 20 mM EDTA in PBS, pH 7.5. The slides were kept in this buffer for up to three days before permanent mounting of the selected slides.

Some of the slides were passed quickly (30 seconds each) through the ethanol series and then through xylene: ethanol mixtures 1:3, 1:1, and 3:1, followed by final 100% xylene before being mounted in Eukitt (Agea Scientific, UK).

A drop of mounting resin was applied to each slide before the section could dry from the xylene. Cover slips were put over the sections with gentle pressure in order to squeeze out the excess mounting media. The slides were left to dry in a fume hood for 3 to 4 days. The slides were cleaned from the mounting resin by soaking in 100% xylene for 10 to 15 minutes or until all the excess resin dissolved into the xylene bath.

5.2.4 Microscopy and photography

The hybridised sections were observed under 4, 10, 20, and 40-x magnification on a Nikon Labophot-2 model microscope in brightfield view.

Photographs of the *in situ* hybridisations were taken before the slides were mounted, and in some cases more pictures were made afterwards from the permanent slides.

A Nikon AF700 camera was used for the pictures. It was attached to the microscope by a 2.5 x adapter. In all cases a Kodak 64 Tungsten Ektackrome film was used.

5.3. Results

The level and pattern of expression of *crk1* gene was investigated in sugar beet tissues from different origin and age by *in situ* hybridisation (EMBO Plant In Situ Hybridisation Instruction Manual, Wageningen, 1997). RNA probes were prepared from cDNA for *crk1* and *Arabidopsis* histone H4 and labelled with digoxigenin.

A relatively uniform distribution of signal was observed for the *crk1* transcripts in the meristematic region of the primary root, as well in emerging lateral root and shoot meristem. Up the root axis, in the elongation zone of the root, the signals were restricted to the central cylinder, the endoderm and epidermis. Further up, at the beginning of the differentiation zone, expression was noted in the central cylinder where the phloem and xylem emerge. A very specific signal was detected in this region and it was only seen in sections where the vascular tissue was just noticeable. When mature vascular bundles formed, this signal was no longer present. The *crk1* expression in primary root sections is shown in figure 5.1.

Figure 5.2 gives a montage picture of the outer layers of 3 month old tap root with the last two cambial layers of cells visible and giving signal with *crk1*. The strongly hybridising areas represent a lateral root primordium (at a 45 degree angle to the section) and a transverse section through other lateral roots. The signal pattern in the lateral roots at the surface of the tap root strongly resembles the one detected in primary roots.

Further, *crk1* transcript signal was detected in the leaf apical meristem, and in the cambial layer of the vascular bundles of the leaf petiole and, also (though to a lesser extent) in the epidermal cells of the petiole.

These results suggest that *Bvcrk1* is expressed in the dividing region of the primary root, in the central cylinder and around the emerging vascular tissue. It is also expressed in the same regions of the lateral roots, and in the numerous cambial layers of the growing tap root. In the tap root, the outer cambium layers seemed more strongly stained than the ones positioned towards the centre of the root. Compared with the expression of the *Arabidopsis* histone H4 in the tap root (pictures not included because of their poor quality), where signal was detected in the last two cambium layers and only in some of the cells, the *Bvcrk1* transcript appeared more uniformly and highly expressed and in the majority of cambium cells. However, when compared to the cdc2 *(Betvu;CDKA;1)*, with expression also detectable within the cambial layers but with much more intense signal, it (figure 5.7), the expression of *Bvcrk1* in similar tissue (figure 5.2 and 5.3) seems weaker. Thus, the signals from the cdc2 ISH come from cells which are actively dividing, and from cells that retain the competence for division. On the

other hand, the Histone H4 is only detectable in cells that are in the G1 and S phases. The *Bvcrk1* expression seems to be constitutive (with no phase specificity) but only in cells undergoing active division, and not like cdc2 in cells with competence to divide.

The uniform signals produced with the *crk1* indicate that this gene is most likely to be expressed throughout the cell cycle. This seems to be supported from previous data about the expression of *Bvcrk1* in cell suspension cultures, where its transcripts were detected after the activation of cell division, but without significant differences between the different cell cycle phases (Fowler *et al.*, 1998).

The fact that the *crk1* message was found to be abundant in root and shoot meristems and showed no organ specificity suggests that the gene may have some role in the basic cell cycle.

It is of interest that meristematic actively dividing tissues on one hand, and in the outer cambium layers of the developing tap roots on the other, present similar activity pattern for the *Bvcrk1* gene.

Of course, the results presented here are not sufficient to give a detailed description of the *crk1* gene activity. However, they provide a basis for more closely targeted and better defined experiments which could help in the understanding of the function of this novel kinase gene and its link to the cell cycle machinery in plant cells.

By further characterisation of the *crk1* gene using expression analyses and especially by studying the protein activity of the gene it can become clear how (and if) this novel protein kinase is related to the cell cycle machinery.

Due to some difficulties in the preparation of sections, a lot of the originally envisaged time points of the sugar beer tap root development were not included in the final *in situ* hybridization experiments. And although this first experiment does not preclude the more detailed study of the tap root in connection with the expression of different cell cycle genes, it indicates that there is no particular reason to go further than the appearance of the first couple of secondary cambium layers. The reason for this is that more efforts could be focused on the detailed description of different cell cycle gene activities within a particular stage of development, for instance the primary root tip. This strategy allows to achieve a more in-depth view about the regulation of sugar beet cell cycle, as numerous genes could be studied. Having a broader picture about several cell cycle genes in the primary root will permit to build up a better experimental set up for further characterisation of the root development.

PAGE NUMBERING AS ORIGINAL

Figure 5.1 Detection of *crk1* mRNA in sugar beet primary root with digoxigenin (DIG)-labelled RNA probe.

Transverse sections of root tip (A), 2-3 mm above the root tip (B), 5 mm above the root tip (C), 2 cm above the root tip (D and E).
A, B, C, and D – hybridisation with antisense *crk1* probe, E – hybridisation with sense *crk1* probe.

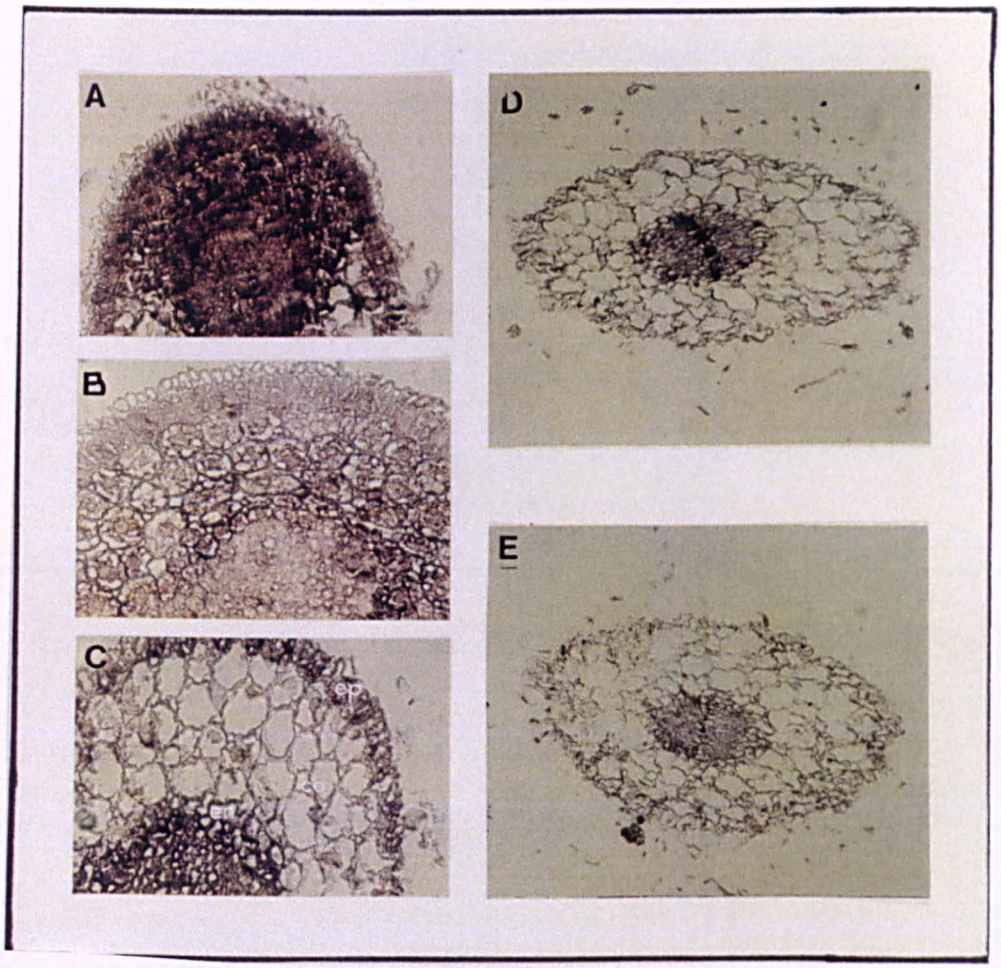


Figure 5.2 Detection of *crk1* mRNA in 3 months old sugar beet root

A and B – Composite pictures of *in situ* hybridisation with *crk1* antisense RNA probe of outer ring transverse sections.
In the picture, the last two cambial rings are visible (A and B). the strong signals in both pictures are localised in the lateral root primordia emerging on the surface of the developing tap root.

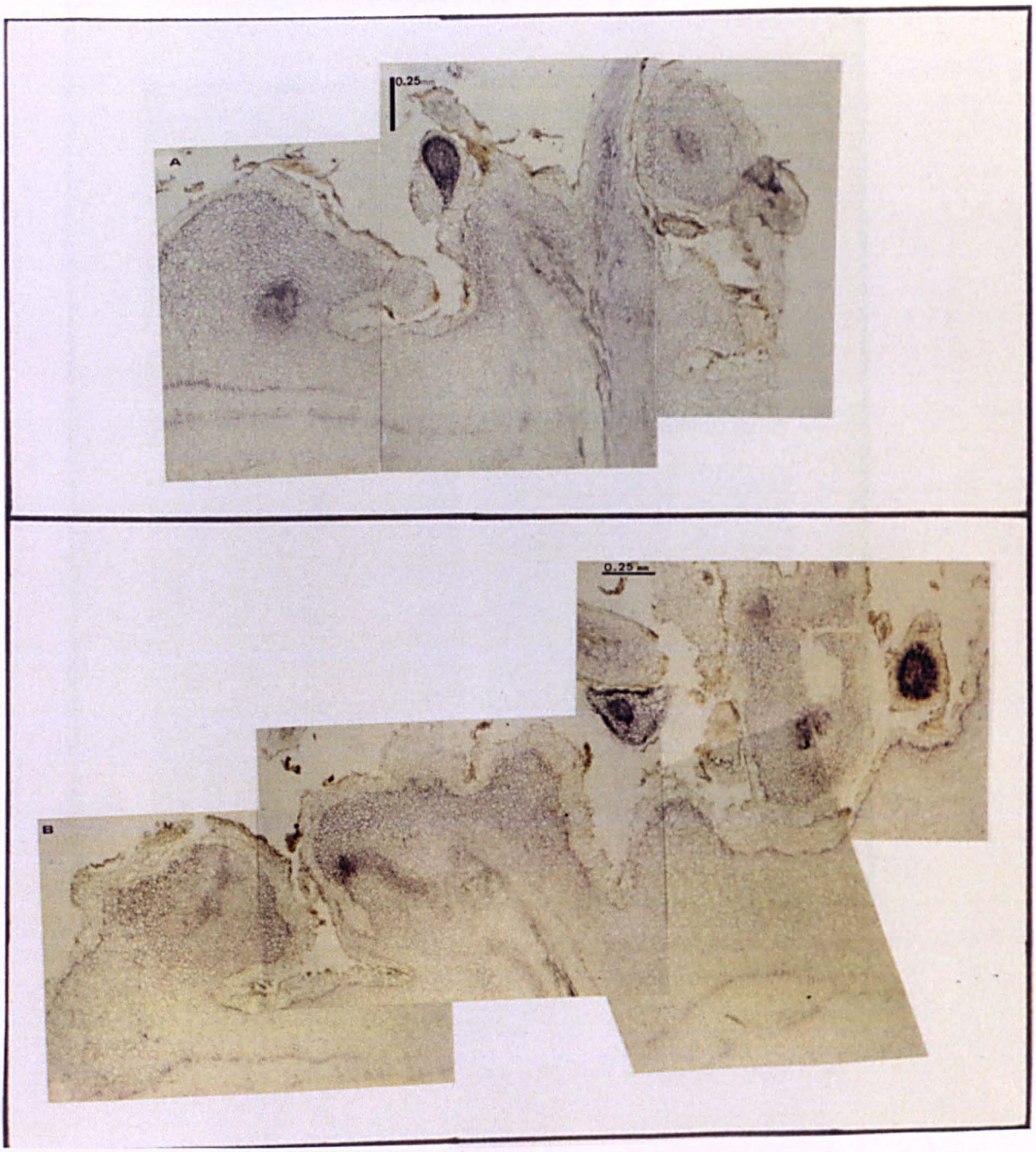


Figure 5.3 Detection of *crk1* mRNA in three months old sugar beet root.

A, B, and D – antisense DIG-labelled probe, C – sense DIG-labelled probe.

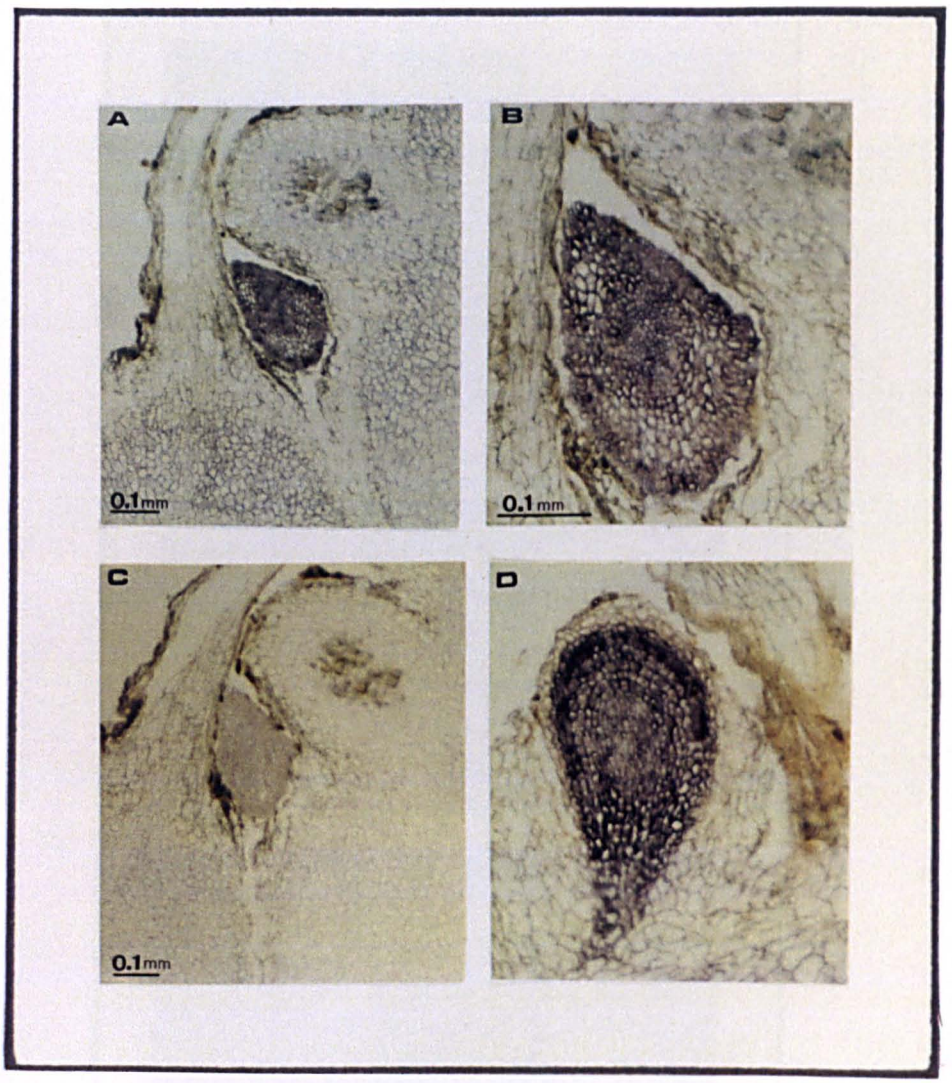


Figure 5.4 **Detection of *crk1* mRNA in transverse sections of shoot apical meristem**

A and B antisense DIG-labelled probe, C – sense DIG-labelled probe. Am –apical meristem, ph – phloem, c – cambium, ep – epidermis.

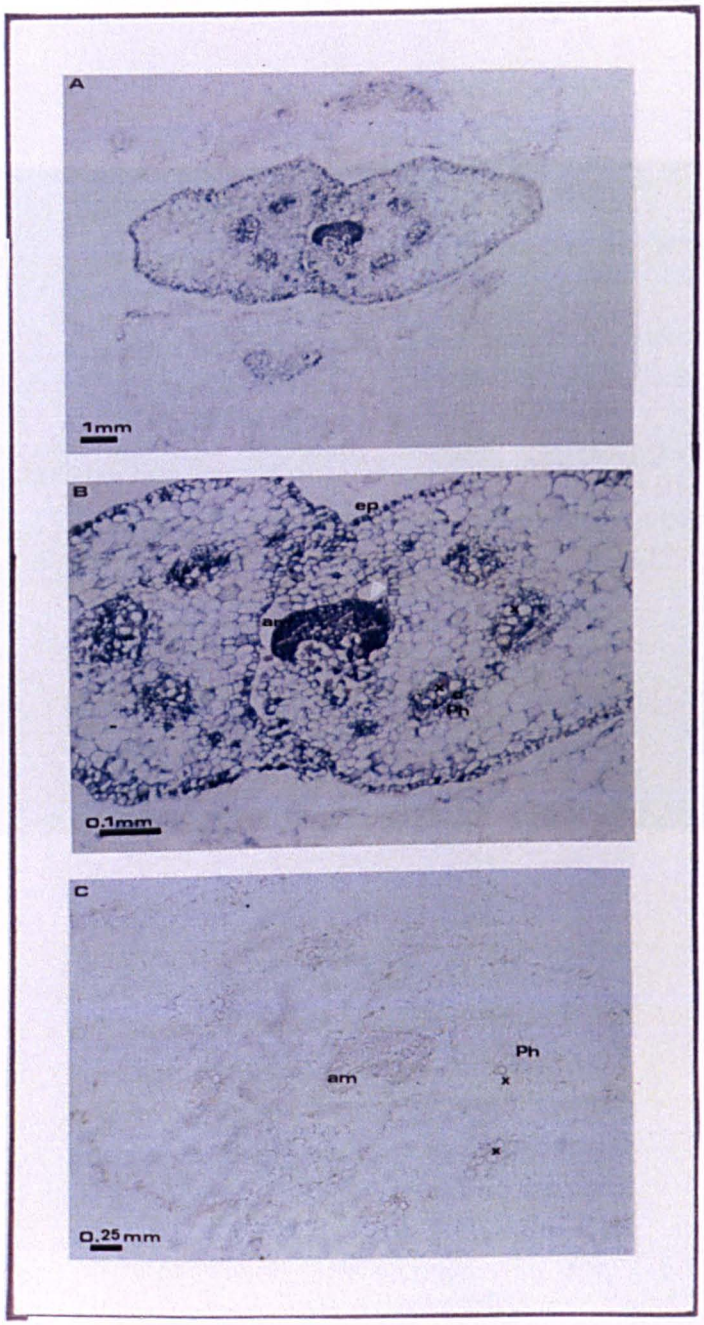
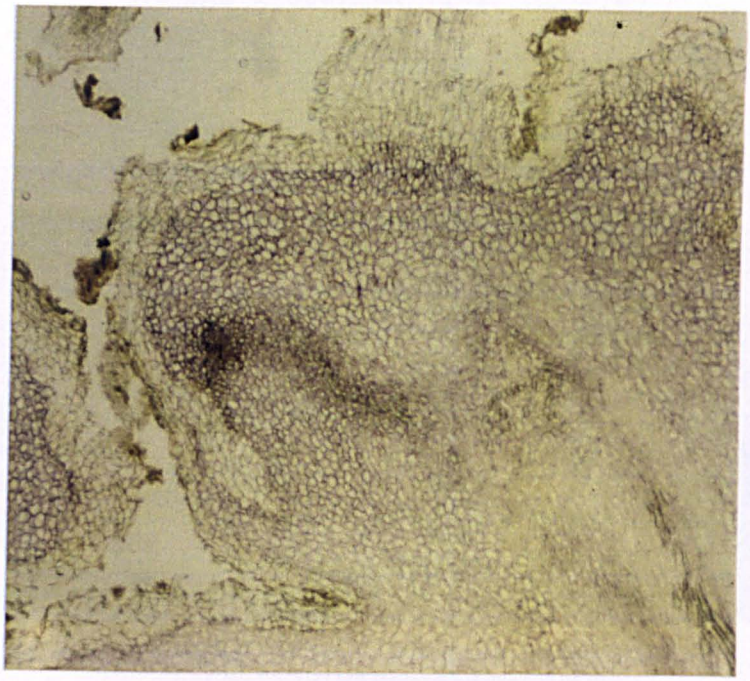


Figure 5.5 Detection of *crk1* mRNA in transverse sections of shoot apical meristem

Dark-field photograph of the section shown in figure 5;4, panel B. The hybridisation signal is seen as pink to purple colour.

Figure 5.6 **Detection of *crk1* mRNA in transverse sections of 3 months old sugar beet root**

A – antisense DIG-labelled probe localisation, B- sense DIG labelled probe localisation.



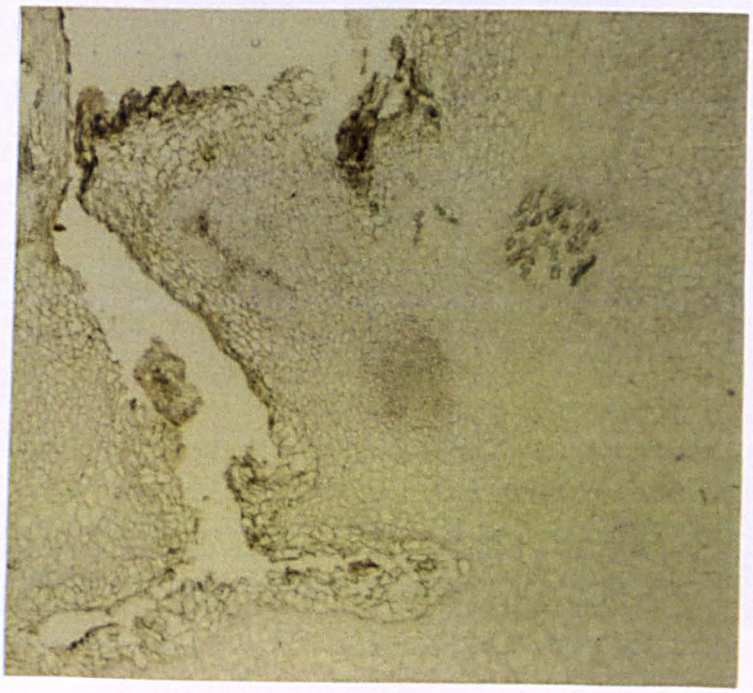
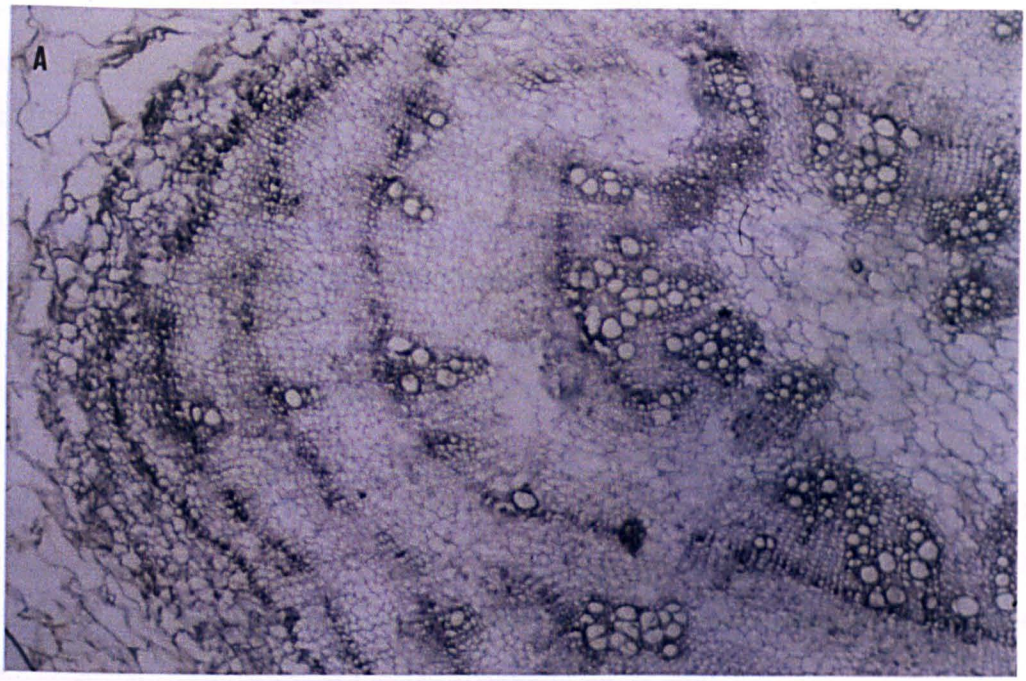
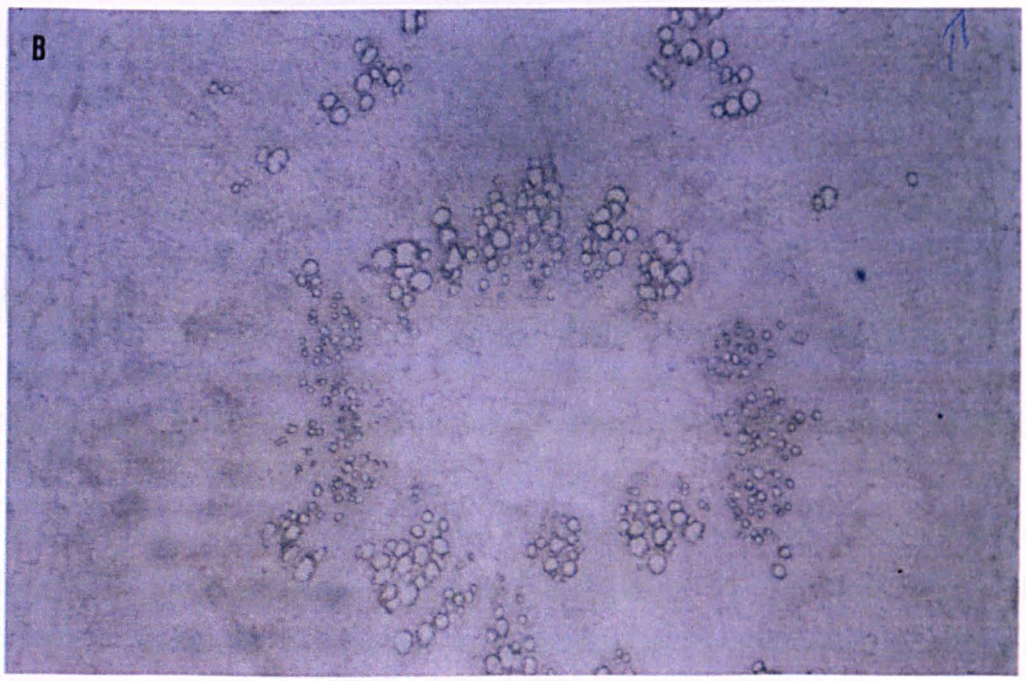


Figure 5.7 Detection of *Bvcdc2* mRNA in 2 months old sugar beet root

A – antisense DIG-labelled probe, B – sense DIG-labelled probe.





PAGE NUMBERING AS ORIGINAL

CHAPTER 6 DISCUSSION

6.1 The plant cell cycle

The plant cell cycle is regulated at two key checkpoints, one in late G1, and the second in late G2 (Huntley and Murray, 1999; Mironov et al., 1999). Progression through the cell cycle depends upon the activities of specific protein kinase complexes, containing a catalytic kinase subunit and a regulatory cyclin subunit. Peaks of kinase activity can be observed at the S phase, and in early M phase, and a drop of activity spans between the two peaks, in G2, and M-G1 transitions (Menges and Murray, 2002). It has been demonstrated that the plant CDK A type kinase is active in both S and G2/M transitions, while the two B-type CDKs are present from the onset of S phase (B1 type) and in G2/M transition (B1 and B2 type). In the last three to four years great progress has been made in our understanding of the regulation of the CDK activities in plant systems. It is now established that D cyclins act as integrators of higher order signals into the cell cycle apparatus, that plants regulate the G1/S transition using the Rb/E2F pathway, and that inhibitors further modulate the activity of CDKs. A great deal of information has been accumulated about the function of the proteolyticaly degradable B cyclins, which are seen as the key CDK regulators for M phase.

As our knowledge about the regulation of the cell cycle expands, more novel components will be added to the regulatory network, with the effect that presently non-annotated genes can be assigned functional characteristics. As new players in the cell cycle machinery are revealed, our attention will fall on the cell cycle in the context of the overall growth and development of the plant. In such terms, the manipulation of the growth and morphological patterns of plants, *via* the manipulation of cell cycle regulators could become a feasible tool for the improvement of crop plants.

6.2 The sugar beet cell cycle and manipulation of storage root development

The "low environmental impact-high yield sugar beet" concept was build up to address some of the problems of the modern high sucrose-yielding sugar beet varieties (Elliott and Weston, 1993; Thomas *et al.*, 1993). In recognition of the high environmental impact of the sugar beet crop, it was hypothesised that genetic manipulations could bring about highly sought changes in the morphology of the storage tap root by achieving a more round and smooth surface. The tap root is build by numerous cambial layers separated by parenchimous tissue. The sugar storage capacity of the root depends on the size of the storage cells and the optimal

concentration of sugar is found in middle-sized cells. Therefore, it was suggested that by manipulating the number of cells, and their size, a higher storage capacity could be reached (Elliott and Weston, 1993). At the same time, it was noted that a root shape which possessed the desired characteristics exists in the table beet cultivar of *Beta vulgaris* but it had reduced storage capacity for sugar. By activation of the outer cambial layers of the table beet, a larger sugar storage capacity can be achieved (Elliott *et al.*, 1993).

The implementation of such improvements requires that we posses detailed understanding of the regulation of the cell cycle activities in the sugar beet root. Any localised alternations of the growth pattern necessitate the identification of gene regulatory elements with capacity for precise tissue and / or cell targeting.

The identification and functional characterisation of the core cell cycle genes from sugar beet comes as a prerequisite for the future efforts in the manipulation of the morphological characteristics of the plant. While working on the isolation of the CDKs from sugar beet, one novel Cdk-like sequence was identified, that of the *Bvcrk1* gene.

6.3 Isolation and characterisation of a novel cdc2-like gene from sugar beet

The isolation and initial characterisation of the *Bvcrk1* gene, described here, form part of the efforts to characterise the genes involved in the sugar beet cell cycle regulation.

A partial *Bvcrk1* sequence was picked out using a PCR RACE approach with degenerate primers designed to the conserved regions of the kinase domain of cdc2 (Kirby, 1996). The partial *Bvcrk1* sequence was used as a probe to screen a sugar beet genomic library prepared for this occasion in the LAMBDA FIX II vector system. A total of $3x10^4$ recombinants were screened and 6 positives were isolated. PCR and restriction digest analysis determined that two of the six clones were highly homologous to the *crk1* probe used for screening. After sequencing it was confirmed that the two clones, designated 7.1 and 8.4, match the *Bvcrk1* gene and represent two alleles of the same gene.

The information obtained from the genomic clones was not sufficient to determine the initiation start site of the genes. This problem was solved by performing a modified RACE – cRACE (Maruyama *et al*, 1995) on sugar beet cDNA isolated from seedlings. This allowed us to identify the transcription start site as well as the 5′-ULS up to the mRNA cap site of the *Bvcrk1* gene. The *Bvcrk1* gene spans 7 exons, and one extra non-translated exon at the 5′-terminus of the gene. Compared to the CDKs, the *Bvcrk1* gene contains an extra exon which carries a domain of unidentified function.

6.4 The predicted Bvcrk1 protein

The full-length *Bvcrk1* gene contains an open reading frame of 1797 bp predicted to encode a protein of 599 amino acids. The predicted Bvcrk1 protein sequence was used for homology searches that revealed that the *Bvcrk1* gene is closely related to the cyclin-dependent protein kinases (Hanks et al., 1988). Apart from the evident kinase domain, the gene contains 5'- and 3'-terminal extensions, which bear no sequence homology to any functionally characterised proteins. The *Bvcrk1* gene has several interesting sequence features. First, although highly homologous in its kinase domain to the CDK family, it lacks a specific Ile or Leu residue within the PSTAIRE domain (figure 4.17), which is invariably present in the proteins identified as CDKs. This has raised the question whether it could interact with cyclins, as the missing residue is thought crucial for the interaction between the kinase and its cyclin partner (Hanks and Hunter, 1995). Second, the gene contains one extra exon at the 5' terminus in comparison to the "PSTAIRE" type kinases. The alignments with the closest plant homologues reveal the presence of a unique N-terminal pattern with no assigned function yet. Further homology searches using only this N-terminal domain demonstrated that this pattern is strictly plant specific. Within the group of *Bvcrk1*-like genes, it is interesting to note the presence of multiple homologues from *Arabidopsis* where 10 genes exhibit sequence and domain organisation similar to the *Bvcrk1* gene. This homology spans the first exon, and includes the altered PSTAIRE region of the type KFMARE. When the *Arabidopsis* family of *Bvcrk1*-like genes is examined, it can be noted that the homology between the family members finishes at the end of the kinase domain, which is followed by divergent extensions of different size. It is still puzzling why such large group of KFMARE-type of CDK-like genes exist and what their individual function might be. It seems likely that the specificity of this group of genes lies outside of the kinase domain, and the N-terminal homologous part.
Two strong nuclear localisation sites were found at the N- and C-terminus of the predicted protein (chapter 5, section 5.5.5), which indicates that the protein is imported to the nucleus and this may represent a way of regulating its activity.

6.5 Possible functions of the *crk1* protein family

The studies of the *in situ* localisation of the *Bvcrk1* transcripts give some idea about the function of the gene. It has been determined that the expression of the gene is confined to the meristematic regions of the plant without any organ specificity. Signals were also detected in the cambial layers of the growing tap root, along the central cylinder and the emerging vascular tissue, the endoderm and epidermis in the primary root. Similar expression patterns

were observed for the shoot apical meristem, with hybridisation signal detected in the meristem, in the cambial layer of the vascular bundles, and to a lesser extent in the epidermal cells of the petiole. The ubiquitous presence of the mRNA in the meristematic tissues and the presence of transcripts in secondary meristems such as the cambium, strongly suggest that the *Bvcrk1* gene has involvement in some aspects of the cell cycle. This correlates with previous data about the expression pattern of the gene in cell suspensions where it was detected after the activation of cell division but without observable differences between the specific cell cycle phases. Later studies with fragments of the promoter of the *Bvcrk1* indicate that the expression of the gene is found in all tissues associated with division in transgenic tobacco seedlings. It is interesting to note that for the pBECKS$_{400}$P-GUS constructs lacking the 5′-UTR (from the translation start site to the ATG start codon), expression was restricted to the shoot apical meristem in the early stages of seedling development (72 hours). By day 5, the difference in GUS expression between constructs either containing or not the 5′ UTR disappeared (A. Slater, personal communication). In transgenic carrot storage organ the expression of the *Bvcrk1* is localised in the cambial layer, while the cdc2a (Arath;CDKA;1) promoter activity was detectable throughout the storage organ (A. Slater, personal communication). The expression of the CDKA;1 is commonly linked to the competence of a given tissue for division, while the genes, which act at more specific timepoints, like for instance the mitotic cyclin B, would produce the so called "patchy" pattern of expression. Interestingly, the expression of the *Bvcrk1* is much more restricted than that of CDKA-type kinase, but unlike other periodical activities, it seems to be present uniformly at transcriptional level in the meristematic tissues. This rises the question whether post translational control has any role in the regulation of the gene activity. Unfortunately, up to date no work has been carried out on the protein of the *Bvcrk1* gene.

6.6 Future work

Further studies will be needed to demonstrate the precise function of the *Bvcrk1* gene. It will be very interesting to try to identify possible interacting partners for this protein, and to see if it can form complexes with any of the known types of plant cyclins. Functional studies of the promoter region of the *Bvcrk1* will also help in pinpointing the exact function of the gene.

In conclusion, it can be said that the involvement of the *Bvcrk1* gene in the regulation of the cell division still needs to be demonstrated directly. Nevertheless, the initial results demonstrate that the gene has some role in the cell cycle and/or early development of the plant cells, and as such clearly presents an interest for further work.

One very intriguing detail is the number of *Bvcrk1* homologues that are present in the *Arabidopsis* genome (already mentioned above). Homology searches show that this type of gene structure is preserved throughout the plant kingdom, as multiple homologues have been reported from rice, maize, and soybean. Information about which protein partners the *Bvcrk1* protein and its homologues have will shed further light for the function of the genes. Knockout mutants of most of the *A. thaliana* homologues to *Bvcrk1* exist and they may serve as an additional source of information for the discovery of the functions of the family.

A further reason for interest in the *Bvcrk1* family comes from the obvious plant specificity of the family. As it is so very often pointed out, plants must have some very specific regulatory features in connection to the fixed way of life, the indeterminate growth and continuous organogenesis, it is very tempting to speculate a link of the *Bvcrk1* gene and its homologues in some of these plant specific aspects in development. Further work with the *Bvcrk1* gene, and undoubtedly, work with the *Arabidopsis* family members will answer this question.

The strongly uniform pattern of expression of the *Bvcrk1* gene in the meristematic regions of developing sugar beet plants, and more specifically, its defined expression in the cambium rings, may just be of further interest to the project of sugar beet improvement. The promoter of the *Bvcrk1* gene is of particular interest, given the strong specificity for actively dividing tissues, and may be used for the targeted regulation of the cell cycle activity in the meristems. A cell cycle specific promoter operating independently of the cell cycle phase, may prove to be a valuable tool for studies of the activities of different cell cycle genes, and a good system for localised gene expression in comparison to constitutively expressed transgenes, or even certain inducible systems.

Finally, the work on the *Bvcrk1* gene characterisation has opened up more questions that it has answered, and it is expected that it has created enough evidence to convince that further research in this field may prove to be very interesting and rewarding.

APPENDIX

Media and solutions

SSC, 20 x:	NaCl	175.3 g	pH 7.0 (NaOH)
	Sodium citrate	88.2 g	
	H_20 to final volume	1 litre	

Denhardt's solution, 50 x:

Ficoll	5g	
Polyvinilpyrroidon	5 g	
BSA, Sigma fraction V	5 g	
H_20 to final volume	500 ml	

LB medium:	NaCl	10 g	pH 7.5(NaOH)
	yeast extract	5 g	
	tryptone	10 g	
	H_20 to final volume	1 litre	

LB agar:	LB medium	pH 7.5 (NaOH)
	1.5 % agar	

LB top agar:	LB medium	
	0.7% agarose	

LB-TAXI:	LB supplemented with	25 µg/ml tetracycline
		50 µg/ml ampicillin
		1 mM IPTG
		40 µg/ml x-GAL

LB-IK:	LB supplemented with	25 µg/ml kanamycin
		1 mM IPTG

NZY medium:	NaCl	5 g	pH 7.5 (NaOH)
	yeast extract	5 g	
	casein hydrolyzate	10 g	
	$MgSO_4.7\ H_20$	2 g	
	H_20 to final volume	1 litre	

PIPES, 10 X	4.0M NaCl
	0.1M PIPES (pH 6.5)

SM Buffer	5.8 g NaCl
	2.0 g $MgSO_4.7\ H_20$
	50 ml 1M Tris-HCl (pH 7.5)
	5.0 ml 2% (w/v) gelatin

Prehybridization and Hybridization solution
for the genomic library screening

2 X PIPES
50% deionized Formamide
5 X Denhardt's
denatured, sonicated salmon sperm DNA (100 µg/µl)

LIST OF PRIMERS USED FOR THE ISOLATION AND SEQUENCING OF THE Bvcrk1 cDNA

primer	length	T_m	primer sequence 5′ to 3′	
P 30	21	69.8	TGG.GCA.AGG.AAC.GTA.TAG.TAA	exon 2
P1 new	22	67.4	TCC.CAG.TTA.ATG.AAT.CTC.GGC.T	exon 2
P8 new	20	63.8	GAA.GCA.TCC.CAT.GAC.AAG.TC	exon 3
P1A 5′	21/9	63	CTC.TAG.CCA.TGA.ACT.TAA.CAC	exon 2
P2A 3′	20/10	64	GCA.CCA.ACT.TAT.CTC.TGG.AC	exon 3
P1B 5′	20/12	68	GAC.TGG.CAG.CAA.GTC.CAG.CT	exon 2
P2B 3′	20/12	68	ACA.ATC.GTG.GGG.TCC.TCC.AC	exon 3
P anchor	20/11	66	AGT.CAC.GTG.ACT.GGA.CAC.AG	
P T_{20} anchor		n/a	AGT.CAC.GTG.ACT.GGA.CAC.AG$(T)_{21}$	
P21 5′P	18	67	P-CCA.GCC.AAT.AAC.TCA.GCC	exon 3
P1 end 5′P	22	69	P-TCC.CAG.TTA.ATG.AAT.CTC.GGC.T	exon 2
P1-cDNA 5′P	22	66	P-CAT.TAA.TTG.ACC.ATC.TTT.CTG.G	5'-ULS
P3′-cDNA	20	67.1	AGT.TGT.ATA.GGG.GTC.CCA.TT	5'-ULS
P5′-cDNA	20	68.5	AGA.AGG.GTT.GAT.GGT.AGC.TG	5'-ULS
P3′-A^{MF}	18/12		GAA.GGC.GAA.CGA.AGG.GCG	exon 1

LIST OF PRIMERS USED FOR THE SEQUENCING OF THE GENOMIC CLONES OF Bvcrk1 GENE

primer	length	Tm	sequence 5′ - 3′	position
P1	21/11	66	CCC.AGT.TAA.TGA.ATC.TCG.GCT	exon 2
P2	20/10	64.7	GTG.GCT.TCT.ATA.CAC.TAG.TC	intron 1
P3	19/9	62.7	GGT.TTC.ATC.CCA.CCT.TAT.C	intron 1
P8	21/9	63.2	AGC.TGA.TTT.TGG.ATT.GGC.TAC	exon 3
P12	20/14	71.5	CAG.GTG.GCT.GCT.GGA.TGG.CC	exon 1
P14	18/10	64.4	GTG.GTT.CGC.CTT.CTG.ATG	exon 4
P17	18/8	60.5	GTT.TGT.GAT.TGC.TGC.TGA	intron 1
P18	17/8	60.2	GAG.AGA.CCT.TTA.GAG.ACT	exon 4
P21	17/10	64	GGC.TGA.GTT.ATT.GGC.TG	exon 3
P22	17/11	66	ATG.CCT.GGC.CGT.ACT.GAG	exon 4
P23	18/14	72	CCC.TGC.TTC.TGC.CGG.GCC	exon7
P24	17/9	62	CCA.GCA.AAG.CAC.AAG.GT	exon 6
P25	17/7	60	TCG.AGT.TT.CTT.TGT.CG	exon 6
Px	20/10	64.7	GAA.CCT.CTT.GCT.TGT.GAT.CC	exon 5
P5′A	23/7		ACT.GAA.ACA.ACA.ACA.ATG.TGA.AA	5'-ULS intron
P5′B	18/7	58.7	AGC.TGG.TGA.AAA.TCT.ATG	5'- ULS
P5′C	19/10	60	CAC.ACT.CAC.CTA.AGT.TGC.G	promoter
P3′D	17/9	62	GAA.CTG.TGT.CAC.GTG.AC	promoter
P3′E	18/6	56.7	TTT.GGA.ATT.AAG.CTT.AGC	intron 4
P5′F	19/11	66.4	CAC.CCA.AGG.GAT.AGC.CAA.G	exon 7
P3′G	18/8	61	GAC.TCT.ACA.TCA.TGA.GCA	promoter
P3′J	19/8	61	CCA.CCT.TTG.CTA.TCT.ATG.A	intron 5
P3′K	17/7	58	ACG.TCA.TTC.ACA.TAC.TC	exon 7
P3′L	17/6	56	ATC.ATG.CAA.TGT.GTG.TT	intron 5
P3′M	18/8	61	CAC.ATT.CAG.CCC.TAG.TTA	3'-UTR
P5′N	18/10	64	GCT.TGG.TGG.ATA.CTT.CGG	exon 5
P3′P	20/7	60	AAT.TGC.TTC.AAG.TGT.CAT.GA	intron 6
P3′Q	17/6	57.9	GCT.TTA.TGC.CTT.TGT.TA	intron 5
P3′R	17/5	50	CTC.ATG.TAT.ATT.TTG.ATC	3'-UTR
P5′S	18/6	53.8	ACA.TAA.ATG.TGT.GAC.ATG	intron 7
P3′T	17/10	56.7	ATG.TGC.TGT.GCC.GAG.TA	3'-UTR
P3′TT	20/6	56.9	TAA.TGT.TAC.ATC.TAA.GGG.TA	3'-UTR
P3′TTT	19/7		CCA.GAT.GAT.ATA.ACA.GGT.T	3'-UTR
P3′W	19/6	57	GAT.ATA.CAT.ATT.GAG.AGG.A	promoter
P3′Y	20/7	62	CTA.AAA.TGT.TGG.GTT.ATC.CAG	intron 6
P3′YY	21/9		ACT.GGA.CAT.CAA.TCT.CAGGAT	intron 6
P3′YYY	18	73.9	TGC.CTA.GAG.CGG.GAG.CTG	intron 6
P3′Z	20/6	60.8	AAG.ATT.ACG.TTT.CCA.ATA.TC	intron 7
P26	20	56.2	CTA.ATA.TTC.ATA.GGA.ACA.GA	intron 6
P-Gap	18	64.9	TTG.TTT.GGG.TGT.GGA.GAT	intron 7
P-intron	20	70.4	GTG.TCA.GGC.CAT.TGC.ATA.AC	intron 7

LIST OF ABBREVIATIONS

ATP	adenosine triphosphate
bp	base pair
cDNA	complementary or copy DNA
CMS	cytoplasmic male sterility
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dH_2O	deionizes water
DNA	deoxyribonucleic acid
dNPT	deoxynucleoside triphosphate
dTTP	deoxythymidine trphosphate
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
gDNA	genomic DNA
GUS	β-gluconidase gene
IPTG	isopropyl-β-D-thiogalactoside
kb	kilobase pair = 1000 bp
LB	Luria-Bretrani
mRNA	messenger ribonucleic acid
OD	optical dencity
PCR	polymerase chain reaction
pfu	phage forming unit
poly(A)RNA	polyadenylated RNA
RNA	ribonucleic acid
SDS	sodium dodecil sulfate
UV	ultra violer
X-Gal	5-bromo,4-chloro,3-indolyl-β-galactopyranoside

NOTE:

For convenience in this text, the non SI measure unit µl is used rather than its SI equivalent $\mathbf{mm^{-3}}$ because in all texts and journals in the area of molecular biology the µl unit is used.

LIST OF REFERENCES

ACH, R.A., P. Taranto, W. Gruissem (1997): A conserved family of WD-40 proteins binds to the retinoblastoma protein in both plants and animals. *Plant Cell* **9**: 1595 - 1606.

ACH, R.A., A.B. Miller, P.C. Zambryski, L; Hanley-Bowdoin, W. Gruissem (1997): RRB1 and RRB2 encode maize retinoblastoma-related proteins that interact with a plant D-type cyclin and geminivirus replication protein. *Mol. Cell. Biol.* **17**: 5077 – 5086.

ADAM, P.D. (2001): Regulation of the retinoblastoma suppressor protein by cyclin/cdks. *BBA* **1471**: M123-M133.

BERGHAMMER, H. and B. Auer (1993): 'Easypreps': Fast and easy plasmid minipreparation for analysis of recombinant clones of *E.coli*. *BioTechniques***14**(4): 527-528.

BERRY, L.D. and K.L. Gould (1996): Regulation of Cdc2 activity by phosphorylation at T14/Y15. *Progress in Cell Cycle Research,* Vol 2, 99-105.

BOLLEN, M., and M. Beullers (2002): Signalling by protein phosphatases in the nucleus. *TRENDS Cell Biol.* **12** (3): 138 – 145.

BONIOTTI M.B., C. Gutierrez (2001): A cell-cycle-regulated kinase activity phosphorylates plant retinoblastoma protein and contains, in Arabidopsis, a CDKA/cyclin D complex. *Plant J* **28**(3):341-50

BOULIKAS, T. (1996): Nuclear import of protein kinases and cyclins. *Journal of Cellular Biochemistry* **60**:61-82.

BROWN, N.R. , M.E.M.Noble, J.A.Endicott, L.N.Johnson (1999):The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nature Cell Biol.*1:438-443.

CHABOUTE, M.E., B. Clement, M. Sekine, G. Philipps, N. Chaubert-Gigot (2000): Cell cycle regulation of the tobacco ribonucleotide reductase small subunit gene is mediated by E2F-like elements. *Plant Cell* **12**:1987-2000.

CHAUBERT-GIGOT, N. (2000): Plant A-type cyclins. *Plant Mol. Biol.***43**: 659-675.

CRIQUI, M., M. Weingrtner, A. Capron, Y. Parmentier, W.H. Shen, E.Heberle-Bors, L. Bogre, P. Genschik (2001): Sub-cellular localisation of GFP-tagged tobacco mitotic cyclins during the cell cycle and after spindle checkpoint activation. *Plant J.* **28**(5): 569-581.

DAHL, M., I. Meskiene, L. B ̄gre, D.T.C. Ha, I. Swoboda, R. Hubman, H. Hirt, E. Heberle-Bors (1995): The D-type alfalfa cyclin gene *cycMs4* complements G1 cyclin-dependent yeast and is induced in the G1 phase of the cell cycle. *Plant Cell* **7**: 1847-1857.

DE BONDT, H.L., J.Rosenblatt, J.Jancarik, H.D.Jones, D.O.Morgan, S-H. Kim (1993): Crystal structure of cyclin -dependant kinase2. *Nature* **363**: 595-602.

de FALCO, G., A. Giordano (1998): CDK9 (PITALRE): a multifunctional cdc2-related kinase. *J Cell Physiol.* **177**(4): 501-506.

de GREGORI, J., G. Leone, A. Miron, L. Jakoi, R.J. Nevis (1997): Distinct roles for E2F proteins in cell growth control and apoptosis. *Proc. Natl; Acad. Sci. USA***94: 7245-7250.**

de JAGER, S.M., M. Menges, U.-M. Bauer, J.A.H. Murray (2001): Arabidopsis E2F1 binds the promoter of S-phase regulated gene AtCDC6 and is a member of a multigene family. *Plant Mol. Biol.***47**:555-568.

DE VEYLDER, L., G. Segers, S, N. Glab, M. Van Montagu, D. Inzé (1997): The Arabidopsis Cks1At protein binds the cyclin-dependent kinases Cdc2aAt and Cdc2bAt. *FEBS Lett.* **412**: 446-452.

DE VEYLDER, L., J. de Almeida-Engler, S; Burssens, A. Manevski, B. Lescure, M. Van Montagu, G. Engler, D. Inzé (1999): A new D-type cyclin of *Arabidopsis thaliana* expressed during lateral root primordia formation. *Planta* **208**: 453-462.

De VEYLDER, L., T. Beeckman, G.T. Beemster, L. Krols, F., Terras, I. Landrieu, E. Van Den Schueren, S. Maes, M. Naudts, D. Inzé (2001): Functional analysis of cyclin-dependent kinase inhibitors in Arabidopsis. *Plant Cell*13: 1653-1668.

DELLA RAGIONE, F., G.L.Russo, A.Olivia, C.Mercurio, S.Mastropietro, V. Della Pietra, V. Zappia (1996): Biochemical characterization of $p16^{INK4}$-and p18-containing complexes in human cell lines. *Journal Biol. Chem.***271**:15942-15949.

DOREÉ, M., S.Galas (1994): The cyclin-dependent protein kinases and the control of cell division. *FASEBS J.* **8**: 1114-1121.

DURFEE, T., H.S. Feiler, W. Gruissem (2000): Retinoblastoma-related proteins in plants: homologues or orthologues of their metazoan counterparts? *Plant Mol. Biol.* **43**:635-642.

EGELKROUT, E.M., D. Robertson, L. Hanley-Broun (2001): Proliferating cell nuclear antigene transcription is repressed through an E2F consensus element and activated by geminivirus infection in mature leaves. *Plant Cell* **13**:1437-1452.

ELLEDGE, S.J. (1996): Cell cycle checkpoints: preventing an identity crisis. *Science* **274**:1664-1672.

ELLIOTT, M.C. and G.D.Weston (1993): *The sugar beet crop: Science into practice*, Chapter 2-Biology and physiology of the sugar-beet plant. Edt D.A.Cooke and R.K.Scott, published by Chapman&Hall, pp 37-66.

ELLIOTT, M.C., A.Slater, M.R.Fowler, J.S.Gartland, M.J.Kirby, N.W.Scott, 1993 BCPC Monograph No 55, Opportunities for molecular biology in crop production New approaches to sugar beet improvement.

EVANS, T., E.T.Rosenthal, J.Youngblom, D. Distel, T. Hunt (1983): *Cell* **33**: 389-396.

EZHZVSKY, S.A., A. Ho, M. Becker-Hapak, P.K. Davis, S.F. Dowdy (2001): Differential regulation of retinoblastoma tumor suppressor protein by G1 Cyclin-dependent kinase complexes in vivo. *Mol. Cell. Biol.* **21**(14): 4773-4784.

FEATHERSTONE C., P. Russell (1991) Fission yeast p107wee1 mitotic inhibitor is a tyrosine/serine kinase. *Nature* **349**(6312): 808-811.

FERREIRA, P., A. Hemerly, R. Villarroel, M. Van Montagu, D. Inzé (1991): The Arabidopsis functional homologue of the $p34^{cdc2}$ protein kinase. *Plant Cell* **3**: 531-540.

FESQUET, D., J.C. Labbe, J. Derancourt, J.P. Capony, S. Galas, F. Girard, T. Lorca, J. Shuttleworth, M. Doree, J.C. Cavadore (1993): The MO15 gene encodes the catalytic subunit of a protein kinase that activates cdc2 and other cyclin-dependent kinases (CDKs) through phosphorylation of Thr161 and its homologues. *EMBO J.* **12**(8): 3111-3321.

FOBERT, P.R., Gaudin V, Lunness P, Coen ES, Doonan JH (1996): Distinct classes of cdc2-related genes are differentially expressed during the cell division cycle in plants. *Plant Cell* **8**(9): 1465-1476.

FOWLER, M.R., M. J. Kirby, N. W. Scott, A. Slater, M. C. Elliott (1998): Induction of cekk division-related genes in quiescent (G0) sugar beet cells. *Physiol. Plantarum* **102**: 61-70.

FRANCIS, D., N.G. Halford (1995): The plant cell cycle. *Physiol.Plant.* **93**: 365-374.

GARRIGA, J., X. Mayol, X, Graña (1996): The CDC2-related kinase PITALRE is the catalytic subunit of active multimeric protein complexes. *Biochem. J.* **319**: 293 - 298.

GAUTIER, J., M.J. Solomon, R.N. Booher, J.F. Bazan, M.W. Krischner (1991):*Cell* **67**: 197–211.

GELSTHORPE, M., M. Pulumati, C. McCallum, K. Dand-Vu, S.I. Tsubota (1997): The putative cell cycle gene, *enhancer of rudimentary*, encodes a highly conserved protein found in plants and animals. *Gene* **186**: 189-195.

GENSCHIK, P., M. Criqui, Y. Parmentier, A. Derevier, J. Fleck (1998): Cell cycle-dependent proteolysis in plants: identification of the destruction box pathway and metaphase arrest produced by the proteasome inhibitor MG132. *Plan Cell* **10**: 2063-2075.

GOULD, K.L., S. Moreno, D.J. Owen, S. Sazer, P. Nurse (1991): *EMBO J.* **10**/ 3297 – 3309.

GRAFI, G., R.J. Burnett, T. Helentjaris, B.A. Larkins, J.A. DECAPRIO? W.R. Sellers, W.G.A. Kaelin (1996): Maize cDAN encoding a mamber of the retinoblastoma protein family – involvement in endoreduplication. *Proc. Natl; Acad. Sci. USA* **93**: 8962 – 8967.

HANKS, S.K. and T.Hunter (1995): The eukaryotic protein kinase superfamily: kinase (catalytic)domain structure and classification. *FASEB J.***9**: 576-596.

HANKS, S.K., A.M.Quinn, T.Hunter (1988): The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**: 42-52.

HASHIMOTO, J., T.Hirabayashi, Y.Hayano, S. Hata, Y. Ohashi, I. Suzuka, T. Utsugi, A.Toh-E, Y.Kikuchi (1992): Isolation and characterization of cDNA clones encoding *cdc2* homologues from *Oryza sativa*: a functional homologue and cognate variants. *Mol.Gen.Genet.* **233**: 10-16.

HEINCHAM, K.A., J.M.Roberts (1994): Rules to replicate by. *Cell* **79**: 557-562.

HEMERLY, A., C.Bergounioux, M. Van Montagu, D.Inzé, P. Ferreira (1992): Genes regulating the plant cell cycle: Isolation of mitotic-like cyclin from *Arabidopsis thaliana. Proc.Natl.Acad.Sci. USA* **89**: 3295-3299.

HEMERLY, A.S., P.Ferreira, J.de Almeida Engler, M.Van Montagu, G.Engler, D.Inzϑ (1993): Cdc2a expression in Arabidopsis is linked with competence for cell division. *Plant Cell* **5**: 1711-1723.

HIRT, H., A.P<y, L.Bögre, I.Meskiene, E. Heberle-Bors (1993): cdc2*MsB*, a cognate *cdc2* gene from alfalfa, complements the G1/S but not the G2/M transition of budding yeast *cdc28* mutants. *Plant Cell* 4(1): 61-69.

HUGHES, M.A. (1996): Plant Molecular Genetics. *Longman, Harlow* pp 10.

HUNTLEY, R., S. Healy, D. Freeman, P. Lavender, S. de Jager, J. Greenwood, J. Makker, E. Walker, M. Jackman, Q. Xie, A. J. Bannister, T. Kouzarides, C. Gutierrez, J.H. Doonan, J.A.H. Murray (1998): The maize retinoblastoma protein homologue ZmRb-1 is regulated during leaf development and displays conserved interactions with G1/S regulators and plant cyclin D (CycD) proteins. *Plant Mol. Biol.* **37**: 155-169.

IMAJUKU, Y., T.Hirayama, H.Endoh, A.Oka (1992): Exon-intron organization of the *Arabidopsis thaliana* protein kinase genes *CDC2a* and *CDC2b*. *FEBS* 1114 **304**(1): 73-77.

ITO, M. (2000): Factors controlling cyclin B expression. *Plant Mol. Biol.***43**: 677-690

ITO, M., S. Araki, S. Matsunaga, T. Itoh, R. Nishihama, Y. Machida, J. H. Doonan, A. Watanabe (2001): G2/M phase-specific transcriiption during the plant cell cycle is mediated by c-Myb-like transcription factors. *Plant Cell* **13**: 1891–1905.

JACOBS, T. (1992): Control of the cell cycle. *Devel.Biol.* **153**: 1-15.

JEFFREY, P.D., A.A. Russo, K. Polyak, E. Gibbs, J. Hurwitz, J. Massague, N.P. Pavletich (1995):Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature***376**:313-320.

JEFFREY, P.D., L. Tong, N.P. Pavletich (2000): structural basis of inhibitionn of CDK-cyclin complexes by INK4 inhibitors. *Genes&Dev.* **14**:3115-3125.

JOUBES, J., C. Chevalier, D. Dudits, E. Heberle-Bors, D. Inzé, M. Umeda, J-P. Renaudin (2000): CDK-related kinases in plants. *Plant Mol.Biol.* **43**: 607-620.

JOUBES, J., M. Lemaire-Chamley, F. Delmas, J; Walter, M. Hernould, A. Mouras, P. Raymond, C.Chevalier, (2001): A new type Cyclin-dependent kinase from tomato expressed in dividing tissues does not interact with mitotic and G1 cyclins. *Plant Physiol.* **126**: 1403-1415.

KATO, J.-Y., M. Matsuoka, D.K. Storm, C.J. Sherr (1994): Regulation of cyclin D-dependant kinase 4 (*cdk4*) by *cdk4*-activating kinase. *Mol.Cell.Biol.* **14**(4): 2713-2721.

KIRBY, M. J. (1995): PhD Thesis, De Montfort University, Leicester.

KIRBY, M.J., M. R. Fowler, N. W. Scott, A. Slater, M. C. Elliott (1998): Cell division-related expression of a novel *cdc2*-like protein kinase gene from sugar beet. Unpublished.

KNIGHTON, D.R., J.H. Zheng, L.F. Ten Eyck, V.A. Ashford, N.H. Xuong, S.S. Taylor, J.M. Sowadski (1991): Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253** (5018): 407-14.

KONG, M., Barnes E.A., Ollendorff V., Donoghue D.J. (2000): Cyclin F regulates the nuclear localization of cyclin B1 through a cyclin-cyclin interaction. *EMBO J* **19**(6):1378-88.

KOSUGI, S., Y. Ohashi (2002): Interaction of the Arabidopsis E2F and DP proteins confers their concomitant nuclear translocation and transactivation. *Plant Physiol.* **128**:833-843.

KOSUGI, S., Y. Ohashi (2002): E2Ls, E2F-like repressors of Arabidopsis that bind to E2F sites in a monomeric form. *J. Biol. Chem.* **277**:16553-16558

LEE, K.-Y., C. C. Helbing, K.-C. Choi, R. N. Johnston, J. H. Wang (1997): Neuronal Cdc2-like kinase (Nclk) binds and phosphorylates the Retinoblastoma protein. *J. Biol. Chem.***272** (9): 5622-5626.

LEE, T.H., M.W. Krischner (1996): An inhibitor of $p34^{cdc2}$ / cyclin B that regulates the G_2 / M transition in *Xenopus* extracts. *Proc. Natl. Acad. Sci. USA* **93**: 352 - 356.

LEES, E.M., E. Harlow (1993): Cyclin: A protein specified by maternal mRNA in sea unchin eggs that is destroyed at each cleavage division.. *Mol.Cell.Biol.* **13**:1194-1201.

LESSARD, P, J.P. Bouly, S. Jouannic, M. Kreis, M. Thomas (1999): Identification of cdc2cAt: a new cyclin-dependent kinase expressed in *Arabidopsis thaliana* flowers. *Biochim Biophys Acta* **1445**(3):351-358.

LUI, H., H. Wang, C. Delong, L.C. Fowke, W.L. Crosby, P.R. Fobert (2000): Cdc2a-interacting protein ICK2 is structurally related to ICK1 and is a potent inhibitor of cyclin-dependent kinase activity in vitro. *Plant J* **21**(4):379-385.

MAGYAR, Z., A. Atanassova, L. DeVeylder, S. Rombauts, D. Inzé (2000): Characterisation of two distinct DP-related genes from Arabidopsis thaliana. *FEBS Lett.***486**: 79-87

MAGYAR, Z., L.Bako, L.B gre, D.Dedeoglu, T.Kapos, D.Dudits (1993): Activated cdc2 genes and cell cycle phase specific cdc2-related kinase complexes in hormone stimulated alfalfa cells. *Plant J.* 4(1): 151-161.

MAGYAR, Z., T. Meszaros, P. Miskolczi, M. Deak, A. Fehér, S. Brown, E. Kondorosi, A. Athanasiadis, S. Pongor, M. Bilgin, L. Bako, C. Koncz, D. Dudits (1997): Cell cycle phase specificty of putative cyclin-dependent kinase variants in synchronised alfalfa cells. *Plant Cell* 9: 223-235.

MARTINEZ, M.C., J.-E.Jørgensen, M.A.Lawton, C.J.Lamb, P.W.Doerner (1992): Spatial pattern of cdc2 expression in relation to meristem activity and cell proliferation during plant development. *Proc.Natl.Acad.Sci.USA* **89**: 7360-7364.

MARUYAMA, I.N., T.L.Rakow , H.I. Maruyama (1995): RACE: a simple method for identification of the 5' end of mRNAs. *Nucleic Acid Research* **23** (18): 3796-3797.

MATSUOKA M., Y. Matsuura, K. Semba, I. Nishimoto (2000): Molecular cloning of a cyclin-like protein associated with cyclin-dependent kinase 3 (cdk 3) in vivo. *Biochem Biophys Res Commun* **273**(2): 442-447.

MATSUSHIME, H., A. M.E. Ewan, D.K. Storm, J.-Y. Kato, S.K. Hanks, C. J. Sherr (1992): Identification and properties of an atypical catalytic subunit (^{p34PSK}-J3/cdk4) for mammalian D type G1 cyclins. *Cell*71: 323-334.

MATSUSHIME. M., D.E. Quelle, S.A. Shurtleff, M. Shibuya, C.J. Sherr, J.Y. Kato (1994): D-type cyclin-dependent kinase activity in mammalian cells. *Mol. Cell. Biol.* **14**: 2066-2076.

McKIBBIN, R.S., N.G. Halford, D. Francis (1998): Expression of fussion yeast *cdc25*alters the frequency of lateral root formation in transgenic tobacco. *Plant Mol. Biol.36: 601-612.*

MENGES, M., and J.A.H. Murray (2002): Synchronous Arabidopsis suspension cultures for analysis of cell-cycle gene activity. *Plant J.* **30**(2): 203-212.

MESKIENE, I., L. Bögre, M. Dahl, M. Pirck, D.T.C. Ha, I. Swoboda, E. Heberle-Bors, G. Ammerer, H. Hirt (1995): *cycMs3*, a novel B-type alfalfa cyclin gene, is isolated in the G_0-to-G_1 transition of the cell cycle. *PlantCell* **7**: 759-771.

MÉSZÁROS, T., P. Miskolczi, F. Ayaydin, A. Pettkó-Szandtner, A. Peres, Z. Magyar, G.V. Horvùath, L. Bakó, A. Fehér, D. Dudits (2000): Multiple cyclin-dependent kinase complexes and phosphatases control G2/M progression in alfalfa cells. *Plant Mol Biol.* **43**:595-605.

MIAO, G.-H., Z.vHong, D.P.S.vVerma (1993): Two functional soybean genes encoding $p34^{cdc2}$ protein kinase are regulated by different plant developmental pathways. *Proc.Natl.Acad.SciUSA.* **90**: 943 -947.

MIRONOV, V., L.Deveylder, M. Van Montagu, D. Inzé (1999): Cyclin-dependent kinases and the cell division in plants: the nexus. *Plant Cell* **11**: 509-521.

MORGAN, O.D., (1995): Principles of CDK regulation. *Nature* **374:** 131-134.

NAGL, W. (1995): Cdc2-kinases, cyclins, and the switch from proliferation to polyploidization. *Protoplasma* **188**: 143-150.

NAKAGAMI, H., M. Sekine, H. Marakami, A. Shinmyo (1999): Tobacco retinoblastoma-related protein phosphorylated by a distinct cyclin-dependent kinase complex with Cdc2/cyclin D *in vitro. Plant J.* **18**: 243-252.

NASMITH, K. (1996): At the heart of the budding yeast cell cycle. *Trends Genet* **12**: 405-412.

NIGG, E. (1995): Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. *BioEssays* **17(6)**: 471-480.

NISHIMURA, A., M. Morita, Y. Nishimura and Y. Sugino (1993): A rapid and highly efficient method for preparation of competent *Escherichia coli* cells. *Nucl. Acid. Res.* **18** (20): 61.

NORBURY, C., P. Nurse (1992): Animal cell cycles and their control. *Annu. Rev. Biochem.* **61**: 440-470.

NURSE, P. (1990): Universal control mechanism regulating onset of M-phase. *Nature* **344**: 503-508.

NURSE, P., (2000): A long twentieth century of the cell cycle and beyond. *Cell* **100** (1): 71-78.

OAKENFULL, A., C. Riou-Khamlichi, J.A.H. Murray (2002): Plant D-type cyclins and the control of G1 progresson. *Phil. Trans. R. Soc. Lond. B* **357**: 749-760.

PARDEE, A.B. (1989): G1 events and regulation of cell proliferation. *Science* **246**(4930):603-8.

PINES, J. and C.L. Rieder (2001): Re-starting mitosis: a contemporary view of mitotic progression. *Nature CellBiol.***3**: E3- E5.

PINES, J.,(1993): Cyclins and cyclin-dependent kinases: take your partners. *Trends Biochen.Sci.* **18**: 195-197.

PINES, J. (1996): Cyclin from sea urchins to HeLas: marking the human cell cycle. *Biochem. Soc. Transact.***24**:15-32.

PORCCEDU, A., H. Stals, J.-P. Reichheld, G. Segers, L. De Vyelder, R. de Pinho Barroco, P. Casteels, M. Van Montagu, D. Inzé, V. Mironov (2001): A plant specific cyclin dependent kinase is involved in the control of G2/M progression in plants. *J. Biol. Chem. Manuscript M011060200*

RAMIREZ-RARRA, E., Q. Xie, M.B. Boniotti, C. Gutierrez (1999): The cloning of plant E2F, a retinoblastoma-binding protein, reveals unique and conserved features with animal G1/S. *Nucl. acids Res.* **27**: 3527-3533.

RECHSTEINER, M., S.W. Rogers (1996): PEST sequences and regulation by proteolysis. *TIBS* **21**: 267-271.

REED, S.I., E. Bailly, V. Dulic, L. Hengst, D. Reznitzky, J. Slingerland (1994): G1 control in mammalian cells. *J. Cell Sci., Supplement* **18**: 69-73.

REICHHELD, J.-P., T. Vermoux, F. Lardon, M. Van Montagu, D. Inzé (1999): Specific checkpoints regulate plant cell cyle progression in response to oxidative stress. *Plant J.* **17**: 647-656.

REICHHELD, J.-P., N. Chaubet, W. H. Shen, J.-P. Renaudin, C. Gigot (1996): Multiple A-type cyclins express sequentially during the cell cycle in *Nicotiana tabacum* BY2 cells. *Proc. Natl. Acad. Sci. USA* **93**: 13819 - 13824.

REN, B., H. Cam, Y. Takahashi, T. Volkert, J. Terrigni, R.A. Young, B.D. Dynlacht (2002): E2F integrates cell cycle progression with DNA repair, replication, and G2/M checkpoints. *Gen. Dev.***16**: 245-256.

RICHHELD, J-P., S. Sonobe, B. Cl9ment, N. Chaubet, C. Gigot (1995): Cell cycle-regulated histone gene expression in synchronized plant cells. *Plant J.* **7**(2): 245-252.

RIUO-KHAMLICHI, C., R. Huntley, A. Jacqmard, .J.A.H. Murray (1999): Cytokinin activation of Arabidopsis cell division through D-type cyclin. *Science* **283**: 1541 – 1544.

ROUDIER, F., E. Fedorova, J. Gyorgyey, A. Feher, S. Brown, A. Kondorosi, E. Kondorosi (2000): Cell cycle function of a *Medicago sativa* A2-type cyclin interacting with PSTAIRE-type cyclin-dependent kinase and a retinoblastoma protein. *Plant J.* **23**: 73-83.

RUSSO, A.A., P.D. Jeffrey, N.P. Pavletich (1996): Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nature Str. Biol.* **3**: 696-700.

SAUTER, M., (1997): Differential expression of a CAK (cdc2-activation kinase)-like protein kinase, cyclins and cdc2 genes from rice during the cell cycle and in response to gibberellin. *Plant J.* **11**: 181-190.

SERGERE J.C., J.Y. Thuret, G. Le, R. Edgardo, D. Carosella, F. Leteurtre (2000): Human CDK10 gene isoforms. *Biochem Biophys Res Commun* **276**(1): 271-277.

SERGES, G., I. Gadisseur, C. Bergounioux, J. de Almeida Engler, A. Jacqmard, M. Van Montagu, D. Inzé (1996): The *Arabidopsis* cyclin-dependent kinase gene cdc2bAt is preferentially expressed during S and G2 phases of the cell cycle. *Plant* .**10**(4): 601-612.

SETIADY, Y.Y., M. Sekine, T. Yamamoto, H. Kouchi, A. Shinmyo (1997): Expression pattern of tobacco cyclin genes. *Plant Cell Rep.* **16**: 368-372.

SETIADY, Y.Y., M.Sekie, N.Hariguchi, T.Yamamoto, H.Kouchi, A.Shinmyo (1995): Tobacco mitotic cyclins: cloning, characterization, gene expression and functional assay. *Plant J.* **8**(6): 949-957.

SHEN, W-H (2000): The plant cell cycle: G1/S regulation. *Euphitica*

SHERR, C.J. (1993): Mammalian G1 cyclins. *Cell* **73**: 1059-1065.

SHERR, C.J. and J.M. Roberts (1999): CDK inhibitors: Positive and negative regulators of G1-phase progression. *Genes&Dev.* **13**: 1501-1512.

SONI, R., J.P. Carmichael, Z.H. Shah, J.A.H. Murray (1995): A family of cyclin D homologs from plants differentially controlled by growth regulators and containing the conserved retinoblastoma protein interaction motif. *Plant Cell* **7**: 85-103.

SORRELL, D. A., M. Menges, J.M.S. Healy, Y. Deveaux, X. Amano, Y. Su, H. Nakagami, A. Shinmyo, J.H. Doonan, M. Sekine, J.A.H. Murray (2001): Cell cycle regulation of cyclin-dependent kinases in tobacco cultivar bright yellow-2 cells. *Plant Phys* .**126**: 1214-1223.

SORRELL, D.A., B. Combettes, N. Chaubet-Gigot, C. Gigot, J.A.H. Murray (1999): Distinct cyclin D genes show mitotic accumulation or constant levels of transcripts in tobacco Bright Yellow-2 cells. *Plant Physiol.* **119**: 343-351

STALS, H., D. Inzé (2001): When plant cells decide to divide. *TRENDS Plant Sci.* **6**: 359-364.

STALS, H., P. Casteels, M. Van Montagu, D. Inzé (2000): Regulation of cyclin dependent kinases in *Arabidopsis tahliana. Plant Mol. Biol.* **43**: 583-593.

STERN, B., P. Nurse (1996): A quantitative model for the cdc2 control of S phase and mitosis in fission yeast. *Trends genet.* **12**: 345-350.

STEVENS, R., L. Mariconti, P. Rossignol, S. Perennes, R. Sella, C. Bergounioux (2002): Two E2F sites in the Arabidopsis MCM3 promoter have different roles in cell cycle activation and meristematic expression. *J. Biol. Chem.* **277(36)**: 32978-32984.

SUN, Y., B.P. Dilkes, C. Zhang , R.A. Dante, N.P. Carneiro, K.S. Lowe, R. Jung, W.J. Gordon-Kamm, B.A. Larkis (1999): Characterisation of maize (Zea mays L.)

Wee1 and its activity in developing endosperm. *Proc. Nat.Acad. Sci. USA* **96**: 4180-4185.

THOMAS, T.H., K.M.A. Gartland, A. Slater, and M.C. Elliott (1993): Opportunities for manipulation of growth and development. *The Sugar beet crop: Science into practice*, Edited by. D.A. Cooke and R.K. Scott, published by Chapman &Hall, pp 512-549.

UMEDA, M., R.P. Bhalerao, J. Schell, H. Uchimiya, C. Koncz (1998): A distinct cyclin-dependent kinase-activating kinase of Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **95**(9):5021-5026.

UMEDA, M., C. Umedahara, M. Yamaguchi, J. Hashimoto, H. Uchimiya (1999): Differential expression of genes for cyclin-dependent protein kinases in rice plants. *Plant Physiol.* **119**: 31-40

VANDEPOELE, K., J. Raes, L. De Veylder, P. Rouzé, S. Rombauts, D. Inzé (2002): Genome-wide analysis of core cell cycle genes in Arabidopsis. *Plant Cell:***14**: 903-916.

WAERNER-WASHBURNE, M., E. Braun, G.C. Johnston, R.A. Singer (1993): Stationary phase in the yeast *Saccharomyses cerevisiae. Microbiol.Rev.***57**: 383-401.

WANG, H., Q. Qi, P. Schorr, A.J. Cutler, W.L. Crosby, L.C. Folke (1998): ICK1, a cyclin-dependent protein kinase inhibitor from Arabidopsis thaliana interacts with both Cdc2a and cycD3, and its expression is induced by abscisic acis *Plant J.***15**: 501-510.

WANG, H., Y. Zhou, S. Glimer, S.. Whitwill, L.C. Folke (2000): Expression of the plant cyclin-dependent kinase inhibitor ICK1 affects cell division, plant growth and morphology. *Plant J.***24**: 613-623.

WEAVER, R.F., and P.W. Hedrick, (1991): Basic Genietics, Wm. C. Brown Publishers, chapter 10, p 240.

WEINERT, T.A. (1997): A DNA damage checkpoint meets the cell cycle engine. *Science* **277**: 1450-1452.

WU, L., T. Ueda, J. Messing (1995): The formation of mRNA 3'-ends in plants. *Plant J*.**8**(3):323-329.

XIE, Q., A. P. Sanz-Burgos, G. J. Hannon, C. Gutiérrez (1996): Plant cells contain a novel member of the retinoblastoma family of growth regulatory proteins. *EMBO J.* **15** (18): 4900 - 4908.

YAMAGUCHI M., T. Fabian, M. Sauter, R.P. Bhalerao, J. Schrader, G. Sandberg, M. Umeda, H. Uchimiya (2000): Activation of CDK-activating kinase is dependent on interaction with H-type cyclins in plants. *Plant J* **24**(1):11-20.

YAMAGUCHI M., M. Umeda, H. Uchimiya (1998): A rice homolog of Cdk7/MO15 phosphorylates both cyclin-dependent protein kinases and the carboxy-terminal domain of RNA polymerase II. *Plant J.* **16**(5): 613-619.

ZACHARIAE,W., K. Nasmyth (1999): Whose end is destruction: cell division and the anaphase-promoting complex. *Genes&Dev.***13**:2039-2058.

ZENG, Y., K. Somasundaram, W. S. El-Diery (1997): AP1 inhibits cancer cell growth and activated $p21^{WAF1/CIP1}$ expression. *Nature Gen.***15**: 78-83.

ZHANG, K., D.S. Lethman, P.C.L. John (1996): Cytokinin controls the cell cycle at mitosis by stimulating the tyrosine dephosphorylation and activation of $p34^{cdc2}$ –like H1 histone kinase. *Planta* **200**: 2-12.